AN INDEX THEOREM FOR NON PERIODIC SOLUTIONS OF HAMILTONIAN SYSTEMS

PAOLO PICCIONE AND DANIEL V. TAUSK

ABSTRACT. We consider a *Hamiltonian setup* $(\mathcal{M}, \omega, H, \mathfrak{L}, \Gamma, \mathcal{P})$, where $(\mathcal{M}, \omega)$ is a symplectic manifold, $\mathfrak{L}$ is a distribution of Lagrangian subspaces in $\mathcal{M}$, $\mathcal{P}$ a Lagrangian submanifold of $\mathcal{M}$, H is a smooth time dependent Hamiltonian function on $\mathcal{M}$ and $\Gamma : [a, b] \to \mathcal{M}$ is an integral curve of the Hamiltonian flow $\vec{H}$ starting at $\mathcal{P}$. We do not require any convexity property of the Hamiltonian function H. Under the assumption that $\Gamma(b)$ is not $\mathcal{P}$-focal it is introduced the Maslov index $\mathrm{i}_{\mathrm{maslov}}(\Gamma)$ of Γ given in terms of the first relative homology group of the Lagrangian Grassmannian; under generic circumstances $\mathrm{i}_{\mathrm{maslov}}(\Gamma)$ is computed as a sort of *algebraic count* of the $\mathcal{P}$-focal points along Γ. We prove the following version of the Index Theorem: under suitable hypotheses, the Morse index of the Lagrangian action functional restricted to suitable variations of Γ is equal to the sum of $\mathrm{i}_{\mathrm{maslov}}(\Gamma)$ and a *convexity term* of the Hamiltonian H relative to the submanifold $\mathcal{P}$. When the result is applied to the case of the cotangent bundle $\mathcal{M} = TM^*$ of a semi-Riemannian manifold (M, g) and to the geodesic Hamiltonian $H(q, p) = \frac{1}{2} g^{-1}(p, p)$, we obtain a semi-Riemannian version of the celebrated Morse Index Theorem for geodesics with variable endpoints in Riemannian geometry.

1. INTRODUCTION

Our interest in the index theory for solutions of the Hamilton equations in a symplectic manifold was originally motivated by the aim of extending to the case of non positive definite metrics the classical results of the Morse theory for geodesics in Riemannian manifolds (see [20]). Despite this original motivation, the geometric applications of the theory developed are left to the very last part of the article, and most of the results presented in the paper belong indeed to the realm of the theory of systems of ordinary differential equations with coefficients in the Lie algebra $\mathrm{sp}(2n, \mathbb{R})$ of the symplectic group. Such systems will be called *symplectic differential systems*.

In order to motivate the theory presented in this paper, we give a short account of the mathematical history of the problem. The origin of the index theory is to be found in the *Sturm theory* for ordinary differential equations (see for instance [5]). The Sturm oscillation theorem deals with second order differential equations of the form $-(px')' + rx = \lambda x$ where p and r are functions with $p > 0$, and λ is a real parameter. The theorem states that the number of zeroes of a non null solution x of the Sturm equation satisfying $x(a) = 0$ equals the index of the index form $I(x_1, x_2) = \int_a^b [p x_1' x_2' + r x_1 x_2] \, \mathrm{d}t$ defined in space of real valued maps on $[a, b]$ vanishing at the endpoints.

The extension of the results of the Sturmian theory to the case of systems of differential equations is essentially due to Morse, obtaining the celebrated Morse Index Theorem in Riemannian geometry (see for instance [6, 20]). The Morse–Sturm systems for which the

Date: revised version of June 2000.

2000 *Mathematics Subject Classification.* 37J05, 53C22, 53C50, 53D12, 70H05.

The first author is partially sponsored by CNPq, the second author is sponsored by FAPESP.

theorem applies are those of the form $g^{-1}(gv')' = Rv$, where $g(t)$ is a *positive definite* symmetric matrix and $R(t)$ is $g(t)$-symmetric linear operator on $I\!R^n$ for all t. Such systems are obtained, for instance, by considering the Jacobi equation along a geodesic in a Riemannian manifold; the equation is converted into a system of ODE's in $I\!R^n$ by means of a parallel trivialization of the tangent bundle of the manifold along the geodesic. In this situation, the *index form* $I(v, w) = \int_a^b \big[g(v', w') + g(Rv, w)\big]\,\mathrm{d}t$ has finite index in the space of vector valued function on $[a, b]$ vanishing at the endpoints, and it is equal to the number of conjugate points of the system in the interval $]a, b[$. By minor changes, the Morse Index Theorem is also valid in the case of *Lorentzian* metrics g, i.e., metrics having index one, provided that one considers causal, i.e., timelike or lightlike, geodesics and that one restricts the index form I to vector fields that are pointwise orthogonal to the geodesic (see [3]). Subsequent results have extended the Index Theorem to the case of solutions with non fixed initial and/or final endpoint (see [15, 23]) and to the case of ordinary differential operators of even order (see [9]).

When passing to the case of spacelike geodesics in Lorentzian manifolds, or, more in general, to geodesics in semi-Riemannian manifolds endowed with metric tensors of arbitrary index, there is no hope to extend the original formulation of the Morse Index Theorem for several reasons. In first place, the action functional is strongly indefinite, i.e., its second variation, given by the index form I, has *always* infinite index. Moreover, the set of conjugate points along a given geodesic may fail to be discrete, and the Jacobi differential operator is no longer self-adjoint.

A different integer valued geometric invariant, called the *Maslov index*, has been recently introduced in the context of semi-Riemannian geodesics (see [11, 13, 19]).

In the case of Riemannian or causal Lorentzian geodesics, the Maslov index coincides with the *geometric index* of the geodesic, which is the number of conjugate points counted with multiplicity. For metrics of arbitrary index, under generic circumstances, this index is computed as a sort of algebraic count of the conjugate points along the geodesic. It is natural to expect that, as in the case of positive definite metrics (see [8]), the Maslov index should play the role of the geometric index in metrics with arbitrary index, and in this paper we present several arguments to strengthen this idea. Namely, we show that the Maslov index, under certain circumstances, is equal to the index of the restriction of I to a suitable subspace of variational vector fields. This equality gives a generalization of the Morse Index Theorem to general indefinite metrics.

The definition of the Maslov index is obtained by developing an intersection theory for curves in the manifold Λ of all Lagrangians of a symplectic space. A non trivial solution of the symplectic differential system gives a continuous curve in the Lagrangian Grassmannian, and the zeroes of the solution correspond to intersections of this curve with the subvariety $\Lambda_{\geq 1}(L_0)$ of all Lagrangians L which are not complementary to a fixed Lagrangian L_0. The Maslov index $\mathrm{i}_{\mathrm{maslov}}$ of the symplectic differential system is then defined using the first relative homology group of the pair $(\Lambda, \Lambda \setminus \Lambda_{\geq 1}(L_0))$. There is no standard notion of Maslov index in the literature, which is defined differently according to the context. We use the definition given in [2, 8, 13, 19]; we remark that for periodic Hamiltonian systems a different definition is usually adopted (see [10, 18, 25]).

Symplectic systems arise naturally as linearizations of the Hamilton equations. We consider a general Hamiltonian setup, consisting of a symplectic manifold $(\mathcal{M}, \omega)$, a Lagrangian submanifold $\mathcal{P}$ of $\mathcal{M}$, a distribution $\mathfrak{L}$ of Lagrangian spaces in $\mathcal{M}$, a time dependent Hamiltonian function H on $\mathcal{M}$ and a given solution $\Gamma : [a, b] \to \mathcal{M}$ of the Hamilton equations, with $\Gamma(a) \in \mathcal{P}$. We introduce an index form I_Γ associated to these objects; for

instance, if we consider the cotangent bundle $\mathcal{M} = TM^*$ of a manifold M, where the distribution $\mathfrak{L}$ is the vertical bundle of TTM^*, and H is a hyper-regular Hamiltonian, then our index form I_Γ coincides with the second variation of the Lagrangian action functional. If M is endowed with a semi-Riemannian metric g and H is given by $H(q,p) = \frac{1}{2}g^{-1}(p,p)$, then the solution Γ projects onto a geodesic γ in (M,g), and I_Γ becomes the standard index form of semi-Riemannian geodesics.

A genuine physical interest in the Hamiltonian extension of the index theory comes from Mechanics and Optics, both classical and relativistic, where the Hamiltonian formalism appears naturally in many situations. For instance, the authors are currently studying a Hamiltonian formalism for light rays in a relativistic medium (see [22], see also [21] for a complete description of the relativistic ray optics). As the underlying spacetime model it is assumed an arbitrary 4-dimensional Lorentzian manifold, and the medium is described by a non convex Hamiltonian function that typically involves the spacetime metric and a number of tensor fields by which the medium is characterized.

Given a semi-Riemannian geodesic $\gamma : [a,b] \to M$, then by a parallel trivialization of the tangent bundle TM along γ the Jacobi equation becomes a Morse–Sturm system. Similarly, given a Hamiltonian setup $(\mathcal{M}, \omega, H, \mathfrak{L}, \Gamma, \mathcal{P})$, then by using a symplectic referential of $T\mathcal{M}$ along Γ adapted to the distribution $\mathfrak{L}$, we obtain a symplectic differential system, that corresponds to the linearization of the Hamilton equations along Γ. Our index form I_Γ will then correspond to a symmetric bilinear form I in the space of maps $v : [a,b] \to I\!R^n$ satisfying suitable boundary conditions. The main result of this paper is an index theorem (Theorem 2.7.10) that relates the index of a suitable restriction of I (or, equivalently, of I_Γ) with the Maslov index of a symplectic differential system. Such results aims at the developments of an infinite dimensional Morse theory for solutions of Hamiltonian systems, in the spirit of [11], where the authors obtain existence results for geodesics in stationary Lorentzian manifolds.

The Hessian of the Hamiltonian function H with respect to the momenta corresponds to a component B of the coefficient matrix of the symplectic differential system. In the case that the Hamiltonian is convex, i.e., when B is positive definite, such index theorem is essentially the classical Morse Index Theorem for Riemannian geodesics, or equivalently, for Morse–Sturm systems with g positive definite. In the non convex case, the index of I is always infinite; in order to obtain a space $\mathcal{K}$ carrying a finite index, we "factor out" an infinite dimensional space $\mathcal{S}$ from the space of all variations, such that I is negative definite on $\mathcal{S}$. The spaces $\mathcal{K}$ and $\mathcal{S}$ are orthogonal with respect to the index form I. The definition of the spaces $\mathcal{K}$ and $\mathcal{S}$ is based on the choice of a *maximal negative distribution* for B, i.e., a family $\{\mathcal{D}_t\}_{t\in[a,b]}$ of subspaces of $I\!R^n$ such that each $\mathcal{D}_t$ is a maximal negative subspace for $B(t)$. The space $\mathcal{K}$ is essentially the space of variations $v : [a,b] \to I\!R^n$ that are "solutions of the symplectic differential system in the directions of $\mathcal{D}$", while the space $\mathcal{S}$ consists of variations taking values in $\mathcal{D}$. A geometric description of this abstract setup in the context of semi-Riemannian geodesics is given in the examples at the end of Section 3. The intersection $\mathcal{K} \cap \mathcal{S}$ consists of solutions of another symplectic differential system, called the *reduced symplectic system*, that vanish at the endpoints. The main technical hypotheses for our theory are a nondegeneracy assumption on the initial condition and the assumption that the reduced symplectic system has no conjugate points on the entire interval $]a,b]$. An even more general result could be proven if this latter condition is dropped, provided that one takes into account in the thesis of the index theorem also the contribution given by the conjugate points of the reduced symplectic system. Some results in this direction were announced in [24].

There exist in the literature other index theories for Hamiltonian systems, that mostly concern the case of periodic solutions. Theorems (like the one presented in this paper) relating a version of the Maslov index of the solution with the index of the second variation of the Lagrangian action functional were proven only in the case of convex Hamiltonians (see for instance [10, 18]). In [7] and [16] the authors develop a Morse theory for periodic solutions of non convex Hamiltonian systems employing a technique that reduces the problem to finite dimensional Morse theory. The index theorems presented in [16, 7] relate the Maslov index of a periodic solution to the Hessian of suitable finite-dimensional versions of the action functional.

The paper is divided into two sections. In Section 2 we study the theory of symplectic differential systems in $\mathbb{R}^n$ and in Section 3 we discuss the applications to Hamiltonian systems and semi-Riemannian geometry. In order to facilitate the reading of the article, we give a short overview of the contents of each subsection. In subsection 2.1 we introduce the notion of symplectic differential system, with the appropriate initial conditions and we define the basic notions, like focal instants and focal index. In subsection 2.2 we state a criterion that says when a second-order linear differential equation in $\mathbb{R}^n$ can be transformed into a symplectic differential system.

In subsection 2.3 we describe briefly the geometry of the Lagrangian Grassmannian of a symplectic space and we define the *Maslov index* associated to a symplectic differential problem. We show that the Maslov index equals the focal index, under a certain non-degeneracy assumption. In subsection 2.4 we show that this non-degeneracy condition is *generic*, i.e., every symplectic problem can be uniformly approximated by symplectic problems satisfying the condition. In subsection 2.5 we define the index form I associated to a symplectic differential problem and we show that, except for the convex case, this index form has always infinite index. In subsection 2.6 we describe the space $\mathcal{K}$ and we introduce the reduced symplectic system. In subsection 2.7 we compute the index of I in $\mathcal{K}$, proving the index theorem. In subsection 2.8, motivated by the index theorem for geodesics with both endpoints variable in submanifolds (see [15, 23]), we extend the index theorem of subsection 2.7 to the case of symplectic differential problems with boundary conditions at both endpoints of the interval. In subsection 2.9 we obtain an alternative version of the index theorem which involves the index of $-I$, by introducing the notion of *opposite* symplectic system. In subsection 2.10 we define a notion of isomorphism for symplectic differential systems, motivated by the fact that Hamiltonian systems determine symplectic differential systems only up to isomorphism, depending on the choice of the symplectic trivialization along the solution. We prove that all the objects of our theory (focal instants, Maslov index, index form) are invariant by such isomorphisms. In subsection 2.11 we show that every symplectic differential system is isomorphic to a Morse–Sturm system. In subsection 3.1 we describe how to produce a symplectic differential problem from a Hamiltonian setup by using a symplectic trivialization of the tangent bundle of the symplectic manifold along a solution. It is then possible, thanks to the results of subsection 2.10, to define the Maslov index and an index form for a Hamiltonian setup. In subsection 3.2 we restate the index theorem in terms of the Hamiltonian setup. In subsection 3.3 we show that, in the case of a hyper-regular Hamiltonian, the index form coincides with the second variation of the action functional associated to the corresponding Lagrangian. In subsection 3.4 we describe the Hamilton equations and their linearized version in terms of a connection in the base manifold M. Finally, in subsection 3.5 we give a geometric application of the index theorem to semi-Riemannian geometry.

2. Symplectic Differential Systems

Throughout the paper we use the following convention: given finite dimensional real vector spaces V and W, we will denote by W^* the dual space of W and, given a linear map $T : V \to W$, we will identify it with the bilinear map $T : V \times W^* \to I\!R$ given by $T(v, \alpha) = \alpha(T(v))$. From the context, there should be no confusion arising from using the same symbol for a linear map and the corresponding bilinear map. The reader should be warned that, for a fixed choice of basis in V and W, the matrix representations of a linear map T and of the corresponding bilinear map do *not* coincide, but they are the transpose of each other. If $V = W$, we observe that the linear map T is invertible if and only if the corresponding bilinear map is nondegenerate. We denote by $T^* : W^* \to V^*$ the adjoint map $T^*(\alpha) = \alpha \circ T$; observe that, in terms of bilinear maps, one has $T(v, \alpha) = T^*(\alpha, v)$ for all $v \in V$ and $\alpha \in W^*$.

As a consequence of these identifications, we have that, if $B : V \times W \to I\!R$ is a bilinear map, and $T : Z \to V$, $S : Z \to W$ are linear maps on some vector space Z, then the composite linear maps $B \circ T$ and $S^* \circ B$ correspond respectively to the bilinear maps $B(T\,\cdot\,, \cdot\,)$ and $B(\,\cdot\,, S\,\cdot\,)$.

Given Banach spaces E_1 and E_2, we denote by $\mathcal{L}(E_1, E_2)$ the set of all bounded linear operators from E_1 to E_2 and by $\mathrm{B}(E_1, E_2, I\!R)$ the set of all bounded bilinear maps from $E_1 \times E_2$ to $I\!R$. If $E_1 = E_2 = E$, we also set $\mathcal{L}(E) = \mathcal{L}(E, E)$ and $\mathrm{B}(E, I\!R) = \mathrm{B}(E, E, I\!R)$; by $\mathrm{B}_{\mathrm{sym}}(E, I\!R)$ we mean the set of symmetric bounded bilinear maps on E. If $E = H$ is a Hilbert space with a fixed inner product $\langle\cdot,\cdot\rangle$, by Riesz's theorem we can identify $\mathcal{L}(H)$ with $\mathrm{B}(H, I\!R)$ by associating the linear operator T to the bilinear form $\langle T\cdot, \cdot\rangle$; obviously, the bilinear form is symmetric if and only if the linear operator is self-adjoint.

We give some general definitions concerning symmetric bilinear forms for later use. Let V be any real vector space and $B : V \times V \to I\!R$ a symmetric bilinear form. The *negative type number* (or *index*) $n_-(B)$ of B is the possibly infinite number defined by

$$n_-(B) = \sup\big\{\dim(W) : W \text{ subspace of } V \text{ on which } B \text{ is negative definite}\big\}. \tag{2.1}$$

The *positive type number* (or *co-index*) $n_+(B)$ is given by $n_+(B) = n_-(-B)$; if at least one of these two numbers is finite, the *signature* $sgn(B)$ is defined by:

$$sgn(B) = n_+(B) - n_-(B).$$

The *kernel* of B, $\mathrm{Ker}(B)$, is the set of vectors $v \in V$ such that $B(v, w) = 0$ for all $w \in V$; the *degeneracy* $dgn(B)$ of B is the (possibly infinite) dimension of $\mathrm{Ker}(B)$. If V is finite dimensional, then the numbers $n_+(B)$, $n_-(B)$ and $dgn(B)$ are respectively the number of 1's, -1's and 0's in the canonical form of B as given by the Sylvester's Inertia Theorem. In this case, $n_+(B) + n_-(B)$ is equal to the codimension of $\mathrm{Ker}(B)$, and it is also called the *rank* of B, $rk(B)$.

2.1. Basic definitions.

We consider the symplectic space $I\!R^n \oplus I\!R^{n*}$ endowed with the canonical symplectic form ω:

$$\omega((v_1, \alpha_1), (v_2, \alpha_2)) = \alpha_2(v_1) - \alpha_1(v_2). \tag{2.2}$$

We denote by $\mathrm{Sp}(2n, I\!R)$ the Lie group of symplectic transformations of the space $(I\!R^n \oplus I\!R^{n*}, \omega)$ and by $\mathrm{sp}(2n, I\!R)$ its Lie algebra. Recall that an element $X \in \mathrm{sp}(2n, I\!R)$ is a linear map $X : I\!R^n \oplus I\!R^{n*} \to I\!R^n \oplus I\!R^{n*}$ such that $\omega(X\,\cdot\,, \cdot\,)$ is symmetric; in block

matrix form, X is given by:

$$X = \begin{pmatrix} A & B \\ C & -A^* \end{pmatrix}, \tag{2.3}$$

where $A : I\!R^n \to I\!R^n$ is an arbitrary linear map, and $B : I\!R^{n*} \to I\!R^n$, $C : I\!R^n \to I\!R^{n*}$ are symmetric when regarded as bilinear maps.

The main object of our study are the following differential systems:

Definition 2.1.1. A *symplectic differential system* in $I\!R^n$ is a first order linear differential system of the form:

$$\begin{cases} v'(t) = A(t)v(t) + B(t)\alpha(t); \\ \alpha'(t) = C(t)v(t) - A^*(t)\alpha(t), \end{cases} \qquad t \in [a,b],\ v(t) \in I\!R^n,\ \alpha(t) \in I\!R^{n*} \tag{2.4}$$

where the coefficient matrix

$$X(t) = \begin{pmatrix} A(t) & B(t) \\ C(t) & -A^*(t) \end{pmatrix}$$

is a curve in $\mathrm{sp}(2n, I\!R)$ (i.e., $B(t)$ and $C(t)$ are symmetric), with A and B of class C^1, C continuous and $B(t)$ invertible for all t. Some of the results will be proven under the more restrictive assumption that C be a map of class C^1 (see Subsection 2.7), however, for the final results, we will get rid of this extra hypothesis by a perturbation argument. Some results concerning the isomorphisms of symplectic systems will be proven under the assumption that B is a map of class C^2 (Subsection 2.11).

Given a symplectic differential system (2.4), we denote by $\Psi(t)$, $t \in [a,b]$, its *fundamental matrix*, defined by:

$$\Psi(t)(v(a), \alpha(a)) = (v(t), \alpha(t)), \tag{2.5}$$

for all solution $(v(t), \alpha(t))$ of (2.4). The map Ψ can be characterized as the curve in the Lie group of linear isomorphisms of $I\!R^n \oplus I\!R^{n*}$ satisfying:

$$\Psi'(t) = X(t)\,\Psi(t), \quad \Psi(a) = \mathrm{Id}. \tag{2.6}$$

Since $X(t) \in \mathrm{sp}(2n, I\!R)$, then $\Psi(t) \in \mathrm{Sp}(2n, I\!R)$ for all t; this means that for any pair of solutions (v, α) and (w, β) of (2.4) we have:

$$\beta(v) - \alpha(w) = \text{constant}. \tag{2.7}$$

Given any C^1-curve $v : [a,b] \to I\!R^n$, the first equation of (2.4) defines a unique continuous curve $\alpha_v : [a,b] \to I\!R^{n*}$; more explicitly:

$$\alpha_v(t) = B(t)^{-1}\Big(v'(t) - A(t)v(t)\Big). \tag{2.8}$$

For future reference, we remark the following equality:

$$\alpha_{fv} = f' B^{-1}(v) + f\alpha_v, \tag{2.9}$$

for any C^1-function $f : [a,b] \to I\!R$.

Let ℓ_0 be *Lagrangian subspace* of $(I\!R^n \oplus I\!R^{n*}, \omega)$. This means that $\dim(\ell_0) = n$ and ℓ_0 is ω-*isotropic*, i.e., ω vanishes on $\ell_0 \times \ell_0$. We consider the following initial conditions for the system (2.4):

$$(v(a), \alpha(a)) \in \ell_0. \tag{2.10}$$

There exists a bijection between the set of Lagrangian subspaces of $(\mathbb{R}^n \oplus \mathbb{R}^{n*}, \omega)$ and the set of pairs (P, S), where P is a subspace of $\mathbb{R}^n$ and $S : P \times P \to \mathbb{R}$ is a symmetric bilinear form. The bijection is defined by:

$$\ell_0 = \big\{(v, \alpha) : v \in P,\ \alpha|_P + S(v) = 0\big\}. \tag{2.11}$$

We can therefore rewrite the initial conditions (2.10) in terms of (P, S):

$$v(a) \in P, \quad \alpha(a)|_P + S(v(a)) = 0. \tag{2.12}$$

For the theory developed in this paper we will make henceforth the following:

Assumption 2.1.2. *We assume that the bilinear form $B(a)^{-1}$ in $\mathbb{R}^n$ is nondegenerate on P, or, equivalently, that $B(a)$ is non degenerate on the annihilator P^o of P, which is easily seen to be given by:*

$$P^o = \big\{\alpha \in \mathbb{R}^{n*} : (0, \alpha) \in \ell_0\big\}.$$

A solution (v, α) of (2.4) will be called an X*-solution*; if (v, α) in addition satisfies (2.10), it will be called an (X, ℓ_0)-solution. A pair (X, ℓ_0) as described above satisfying Assumption 2.1.2 will be called a *set of data for the symplectic differential problem*. By a slight abuse of terminology, we will say that a C^2-curve $v(t)$ is an X-solution if the pair (v, α_v), with α_v given by (2.8), is an X-solution; similarly, we say that v is an (X, ℓ_0)-solution if the pair (v, α_v) is an (X, ℓ_0)-solution.

We now consider a fixed set of data (X, ℓ_0) for the symplectic differential problem. We denote by $\mathbb{V}$ the set of all (X, ℓ_0)-solutions:

$$\mathbb{V} = \big\{v : v \text{ is an } (X, \ell_0)\text{-solution}\big\}. \tag{2.13}$$

The fact that ℓ_0 is Lagrangian, combined with (2.7) implies that

$$\alpha_v(w) = \alpha_w(v), \quad \forall\, v, w \in \mathbb{V}. \tag{2.14}$$

For $t \in [a, b]$, we set

$$\mathbb{V}[t] = \big\{v(t) : v \in \mathbb{V}\big\}.$$

From (2.14) and a simple dimension counting argument, the annihilator of $\mathbb{V}[t]$ in $\mathbb{R}^{n*}$ is given by:

$$\mathbb{V}[t]^o = \big\{\alpha_v(t) : v \in \mathbb{V},\ v(t) = 0\big\}, \quad t \in [a, b]. \tag{2.15}$$

Definition 2.1.3. An instant $t \in\,]a, b]$ is said to be *focal* for the pair (X, ℓ_0) if there exists a non zero $v \in \mathbb{V}$ such that $v(t) = 0$, i.e., if $\mathbb{V}[t] \neq \mathbb{R}^n$. The *multiplicity mul*$(t)$ of the focal instant t is defined to be the dimension of the space of those $v \in \mathbb{V}$ vanishing at t, or, equivalently, the codimension of $\mathbb{V}[t]$ in $\mathbb{R}^n$. The *signature sgn*(t) of the focal instant t is the signature of the restriction of the bilinear form $B(t)$ to the space $\mathbb{V}[t]^o$, or, equivalently, the signature of the restriction of $B(t)^{-1}$ to the $B(t)^{-1}$-orthogonal complement $\mathbb{V}[t]^\perp$ of $\mathbb{V}[t]$ in $\mathbb{R}^n$. The focal instant t is said to be *nondegenerate* if such restriction is nondegenerate. If there is only a finite number of focal instants in $]\,a, b]$, we define the *focal index* $\mathrm{i}_{\rm foc} = \mathrm{i}_{\rm foc}(X, \ell_0)$ to be the sum:

$$\mathrm{i}_{\rm foc} = \sum_{t \in\,]a,b]} \textit{sgn}(t). \tag{2.16}$$

Remark 2.1.4. There exist several situations where the number of focal instants is indeed finite; for instance, this is always the case when the coefficient matrix X is real analytic in t. This fact follows from the observation that the focal instants are the zeroes of the function $\det(v_1(t), \ldots, v_n(t))$, where $\{v_1, \ldots, v_n\}$ is a basis of $\mathbb{V}$. Such function does not vanish

identically, as we will see in Subsection 2.3 that it does not vanish in a neighborhood of $t = a$. We will prove in Subsection 2.3 that nondegenerate focal instants are isolated.

In a sense, Assumption 2.1.2 says that $t = a$ is a nondegenerate focal instant; observe that $\mathbb{V}[a] = P$.

2.2. **Morse–Sturm and second order linear differential equations.** An important class of examples of symplectic differential systems arises from the so called *Morse–Sturm systems*, which are second order differential equations in $\mathbb{R}^n$ of the form:

$$g^{-1}(gv')' = Rv, \tag{2.17}$$

where $g(t)$ is a C^1-curve of symmetric nondegenerate bilinear forms in $\mathbb{R}^n$ and $R(t)$ is a continuous curve of linear operators on $\mathbb{R}^n$ such that $g(t)(R(t)\cdot, \cdot)$ is symmetric for all t. Considering the change of variable $\alpha = gv'$, the equation (2.17) is equivalent to the following symplectic system:

$$\begin{cases} v' = g^{-1}\alpha, \\ \alpha' = gRv. \end{cases} \tag{2.18}$$

Comparing with (2.4), in (2.18) we have $A = 0$, $B = g^{-1}$, $C = gR$.

In general, (2.4) may be written as a second order equation, as follows. Substitution of (2.8) into the second equation of (2.4) shows that v is a solution of (2.4) if and only if it is a solution of the following second order equation:

$$\left[B^{-1}(v' - Av)\right]' = Cv - A^*B^{-1}(v' - Av). \tag{2.19}$$

One has the following natural question: when is the second order linear equation

$$v'' + Z_1v' + Z_2v = 0, \tag{2.20}$$

with $Z_1, Z_2 : [a, b] \to \mathcal{L}(\mathbb{R}^n)$ continuous, of the form (2.19) for some symplectic differential system? It is not hard to prove the following:

Proposition 2.2.1. *The differential equation* (2.20) *arises from a symplectic system if and only if there exists a (fixed) symplectic form $\overline{\omega}$ in $\mathbb{R}^{2n}$ such that $\Omega(t)^{-1}(\{0\} \oplus \mathbb{R}^n)$ is Lagrangian in $(\mathbb{R}^{2n}, \overline{\omega})$ for all t, where $\Omega(t) : \mathbb{R}^{2n} \to \mathbb{R}^{2n}$ is the fundamental matrix of* (2.20)*.*

Proof. Assume that $\xi(t) = \Omega(t)^{-1}(\{0\} \oplus \mathbb{R}^n)$ is Lagrangian in $(\mathbb{R}^{2n}, \overline{\omega})$ for all t and choose a C^1-family of symplectomorphisms $\phi(t) : (\mathbb{R}^{2n}, \overline{\omega}) \to (\mathbb{R}^n \oplus \mathbb{R}^{n*}, \omega)$ such that $\phi(t)(\xi(t)) = \{0\} \oplus \mathbb{R}^{n*}$ for all t. Define $\Psi(t) = \phi(t)\phi(a)^{-1}$ and $X(t) = \Psi'(t)\Psi(t)^{-1}$; then X is the coefficient matrix of a symplectic system that corresponds to (2.20) by the change of variables $(v(t), \alpha(t)) = \phi(t)\Omega(t)^{-1}(v(t), v'(t))$. □

2.3. **The Maslov Index.** The goal of this subsection is to produce an integer-valued invariant for a set of data (X, ℓ_0) for the symplectic differential problem. We call this number the *Maslov index* of the pair (X, ℓ_0), we show that it is stable by C^0-small perturbations of the data, and that it is *generically* equal to the focal index $\mathrm{i}_{\mathrm{foc}}(X, \ell_0)$. We will use several well known facts about the geometry of the Lagrangian Grassmannian of a symplectic space (see for instance [2, 8, 12, 19, 26]); in particular, we will make full use of the notations and of the results proven in Reference [19].

Recalling the definition of Ψ in formula (2.5), we start by observing that the space:

$$\ell(t) = \Psi(t)(\ell_0) = \big\{(v(t), \alpha_v(t)) : v \in \mathbb{V}\big\} \tag{2.21}$$

is Lagrangian for all t. Let Λ be the set of all Lagrangian subspaces of $(\mathbb{R}^n \oplus \mathbb{R}^{n*}, \omega)$; Λ is a compact, connected real analytic embedded $\frac{1}{2}n(n+1)$-dimensional submanifold of the

Grassmannian of all n-dimensional subspaces of $I\!R^n \oplus I\!R^{n*}$. For $L \in \Lambda$, the tangent space $T_L\Lambda$ can be canonically identified with the space $\mathrm{B}_{\mathrm{sym}}(L, I\!R)$ of all symmetric bilinear forms on L.

Let $L_0 \in \Lambda$ be fixed; we define the following subsets of Λ:

$$\Lambda_k(L_0) = \big\{L \in \Lambda : \dim(L \cap L_0) = k\big\}, \quad k = 0, \dots, n. \tag{2.22}$$

Each $\Lambda_k(L_0)$ is a connected embedded analytic submanifold of Λ having codimension $\frac{1}{2}k(k+1)$ in Λ; $\Lambda_0(L_0)$ is a dense open contractible subset of Λ, while its complementary set:

$$\Lambda_{\geq 1}(L_0) = \bigcup_{k=1}^{n} \Lambda_k(L_0) \tag{2.23}$$

is *not* a regular submanifold of Λ. It is an algebraic variety, and its regular part is given by $\Lambda_1(L_0)$, which is a dense open subset of $\Lambda_{\geq 1}(L_0)$. Observe that $\Lambda_1(L_0)$ has codimension 1 in Λ; moreover, it has a *natural* transverse orientation in Λ, which is canonically associated to the symplectic form ω (see [19, Section 3] for the proofs and the details of these results).

The first relative singular homology group with integer coefficients $H_1(\Lambda, \Lambda_0(L_0))$ is computed in [19, Section 4]:

$$H_1(\Lambda, \Lambda_0(L_0)) \simeq \mathbb{Z}, \tag{2.24}$$

where the choice of the above isomorphism is related to the choice of a transverse orientation of $\Lambda_1(L_0)$ in Λ, and it is therefore canonical. Every continuous curve l in Λ with endpoints in $\Lambda_0(L_0)$ defines an element in $H_1(\Lambda, \Lambda_0(L_0))$, and we denote by

$$\mu_{L_0}(l) \in \mathbb{Z}$$

the integer number corresponding to the homology class of l by the isomorphism (2.24). This number, which is additive by concatenation and invariant by homotopies with endpoints in $\Lambda_0(L_0)$, is to be interpreted as an *intersection number* of the curve l with $\Lambda_{\geq 1}(L_0)$ (see Proposition 2.3.2 below).

Recalling formula (2.21), we note that ℓ is a C^1-curve in Λ; we fix

$$L_0 = \{0\} \oplus I\!R^{n*}$$

and we observe that the focal instants of the pair (X, ℓ_0) coincide with the intersections of the curve ℓ with $\Lambda_{\geq 1}(L_0)$. Moreover, $\ell(t) \in \Lambda_k(L_0)$ if and only if *mul*$(t) = k$.

We want to define the *Maslov index* $\mathrm{i}_{\mathrm{maslov}}(X, \ell_0)$ to be the integer $\mu_{L_0}(\ell)$; however, we observe that the curve ℓ may fail to have endpoints in $\Lambda_0(L_0)$. In order to have $\ell(b) \in \Lambda_0(L_0)$, we need to assume that $t = b$ is not focal. The initial point $\ell(a)$ is not in $\Lambda_0(L_0)$ unless $P = I\!R^n$; however, we will prove below that, because of Assumption 2.1.2, the curve ℓ has at the most an isolated intersection with $\Lambda_{\geq 1}(L_0)$ at $t = a$. We can therefore give the following:

Definition 2.3.1. Assume that $t = b$ is not a (X, ℓ_0)-focal instant. The Maslov index $\mathrm{i}_{\mathrm{maslov}} = \mathrm{i}_{\mathrm{maslov}}(X, \ell_0)$ is defined as:

$$\mathrm{i}_{\mathrm{maslov}} = \mu_{L_0}(\ell|_{[a+\varepsilon, b]}), \tag{2.25}$$

where $\varepsilon > 0$ is small enough so that ℓ does not intercept $\Lambda_{\geq 1}(L_0)$ in $]a, a+\varepsilon]$.

Clearly, the quantity on the right hand side of equality (2.25) does not depend on the choice of ε.

We now describe a method for computing $\mu_{L_0}(l)$ of a C^1-curve l in Λ under some generic conditions. Recall that tangent vectors to Λ are canonically identified with symmetric bilinear forms.

Proposition 2.3.2. *Let $l : [a,b] \to \Lambda$ be a curve of class C^1, and let $t_0 \in [a,b]$ be such that $l(t_0) \in \Lambda_{\geq 1}(L_0)$ and $l'(t_0)$ is non degenerate on $l(t_0) \cap L_0$. Then, t_0 is an isolated intersection of l with $\Lambda_{\geq 1}(L_0)$. Moreover, if l has endpoints in $\Lambda_0(L_0)$ and all the intersections of l with $\Lambda_{\geq 1}(L_0)$ satisfy the above nondegeneracy condition, then $\mu_{L_0}(l)$ can be computed as:*

$$\mu_{L_0}(l) = \sum_{t \in \,]a,b[} \mathit{sgn}(l'(t)|_{l(t)\cap L_0}). \tag{2.26}$$

Proof. See [19, Corollary 4.3.3]. □

Observe in particular that, under the assumptions of Proposition 2.3.2, if all the intersections of l with $\Lambda_{\geq 1}(L_0)$ lie in $\Lambda_1(L_0)$ and are transversal to $\Lambda_1(L_0)$, then, by (2.26), μ_{L_0} is computed as the difference between the number of positively oriented intersections and the number of negatively oriented intersections of l with $\Lambda_1(L_0)$.

Using (2.6) and (2.21), a simple computation yields:

$$\ell'(t) = \omega(X(t)\,\cdot\,,\,\cdot\,)|_{\ell(t)}. \tag{2.27}$$

Theorem 2.3.3. *Let (X, ℓ_0) be a set of data for the symplectic differential problem. For $\varepsilon > 0$ sufficiently small, there are no focal instants in $]a, a+\varepsilon]$. If $t_0 \in]a,b]$ is a nondegenerate focal instant, then t_0 is an* isolated *focal instant. If this nondegeneracy condition is satisfied by all focal instants and if $t = b$ is not focal, then the following equality holds:*

$$\mathrm{i}_{\mathrm{maslov}}(X, \ell_0) = \mathrm{i}_{\mathrm{foc}}(X, \ell_0). \tag{2.28}$$

Proof. For all $t \in [a,b]$, the projection onto the second coordinate gives an identification between $\ell(t) \cap L_0$ and the space $\mathbb{V}[t]^o$ (recall formula (2.15)). From (2.27), it follows easily that this identification carries the restriction of $\ell'(t)$ to the restriction of $B(t)$. The conclusion follows then from Proposition 2.3.2, observing that, by Assumption 2.1.2, $B(a)$ is non degenerate on $P^o = \mathbb{V}[a]^o$. □

Remark 2.3.4. Suppose that we are given a family $\{X_\delta\}$ of coefficient matrices for symplectic differential systems such that $(t,\delta) \mapsto X_\delta(t)$ is continuous and δ runs on a compact topological space. Then, it is not hard to show that we can find $\varepsilon > 0$ independent on δ such that there are no (X_δ, ℓ_0)-focal instants in $]a, a+\varepsilon]$.

Remark 2.3.5. By the homotopical invariance of μ_{L_0} and by Remark 2.3.4, it follows that the Maslov index of (X, ℓ_0) is stable by continuous deformations of the coefficient matrix X, as long as the instant $t = b$ remains non focal during the deformation. In particular, if $t = b$ is not focal, the Maslov index is stable by uniformly small perturbations of X.

2.4. **On the nondegeneracy condition.** We have seen in the previous subsection that the equality between the Maslov and the focal index of a pair (X, ℓ_0) holds under the assumption of nondegeneracy for the (X, ℓ_0)-focal instants. We will call *nondegenerate* a pair (X, ℓ_0) for which all the focal instants are nondegenerate. The aim of this section is to prove that this condition is *generic*, i.e., that every pair (X, ℓ_0) can be uniformly approximated by nondegenerate pairs.

Proposition 2.4.1. *Let (X, ℓ_0) be a set of data for the symplectic differential problem. There exists a sequence $X_k : [a,b] \to \mathrm{sp}(2n, I\!R)$ of smooth curves such that X_k tends to X uniformly as $k \to \infty$, and (X_k, ℓ_0) is a nondegenerate set of data for the symplectic*

differential problem for all k. Moreover, if $t=b$ is not (X,ℓ_0)-focal, then, for k sufficiently large, $t=b$ is not (X_k,ℓ_0)-focal.

Proof. Since X can be uniformly approximated by smooth curves, there is no loss of generality in assuming that X is smooth. The first step is to prove that ℓ can be approximated in the C^1-topology by smooth curves of Lagrangians starting at ℓ_0 having intersections with $\Lambda_{\geq 1}(L_0)$ only in $\Lambda_1(L_0)$ and transversally.

Let $\varepsilon>0$ be such that $\ell|_{]a,a+\varepsilon]}$ does not intersect $\Lambda_{\geq 1}(L_0)$. It follows from [14, 2.1. Transversality Theorem (a)] that we can find a sequence $\ell_k:[a+\varepsilon,b]\to\Lambda$ of smooth curves that are transverse to $\Lambda_r(L_0)$ for all $r=1,\ldots,n$ and such that ℓ_k converges to $\ell|_{[a+\varepsilon,b]}$ in the C^1-topology. Since $\Lambda_r(L_0)$ has codimension greater than one for $r\geq 2$, transversality to $\Lambda_r(L_0)$ in facts implies that ℓ_k does not intercept $\Lambda_r(L_0)$. Hence, ℓ_k has only transverse intersections with $\Lambda_1(L_0)$. Clearly, each ℓ_k can be extended to a smooth curve on $[a,b]$ such that $\ell_k=\ell$ in $[a,a+\frac{\varepsilon}{2}]$ and such that ℓ_k converges to ℓ on $[a,b]$ in the C^1-topology. For k sufficiently large, it will also follow that ℓ_k does not intercept $\Lambda_{\geq 1}(L_0)$ in $]a,a+\varepsilon]$.

It's now possible[1] to construct a sequence $\Psi_k:[a,b]\to\mathrm{Sp}(2n,I\!\!R)$ of smooth curves converging to Ψ in the C^1-topology such that $\Psi_k(a)=\mathrm{Id}$ for all k and such that each Ψ_k projects into ℓ_k by evaluation at ℓ_0.

We now set, for each k, $X_k=\Psi_k^{-1}\Psi_k'$. Obviously X_k is a smooth curve in $\mathrm{sp}(2n,I\!\!R)$ and X_k tends uniformly to X. Moreover, denoting by B_k the upper-right $n\times n$ block of X_k, we see that B_k is invertible for k sufficiently large, since B_k tends uniformly to B and B is invertible. Similarly, since $B(a)$ is nondegenerate on P^o (recall assumption 2.1.2), $B_k(a)$ is nondegenerate on P^o for k sufficiently large. Therefore (X_k,ℓ_0) is a set of data for the symplectic differential problem, and obviously ℓ_k is its associated curve of Lagrangians. It follows that (X_k,ℓ_0) is nondegenerate.

Finally, if $t=b$ is not (X,ℓ_0)-focal then $\ell(b)\in\Lambda_0(L_0)$. Since $\Lambda_0(L_0)$ is open and $\ell_k(b)$ tends to $\ell(b)$ we see that $t=b$ is not (X_k,ℓ_0)-focal for k sufficiently large. $\square$

We conclude with the observation that, by Remark 2.3.5, in Proposition 2.4.1 we have:

$$\mathrm{i}_{\mathrm{maslov}}(X_k,\ell_0)=\mathrm{i}_{\mathrm{maslov}}(X,\ell_0),$$

for k sufficiently large.

2.5. **The index form.** We denote by $L^2([a,b];I\!\!R^n)$ the Hilbert space of square integrable $I\!\!R^n$-valued functions on $[a,b]$, and, for $j\geq 1$, by $H^j([a,b];I\!\!R^n)$ the Sobolev space of functions of class C^{j-1}, with $(j-1)$-th derivative absolutely continuous and with square integrable j-th derivative. We also denote by $H^1_0([a,b];I\!\!R^n)$ the subspace of $H^1([a,b];I\!\!R^n)$ consisting of those functions vanishing at a and b; for a given subspace $P\subset I\!\!R^n$ let $H^1_P([a,b];I\!\!R^n)$ be the subspace of $H^1([a,b];I\!\!R^n)$ consisting of those maps v such that $v(a)\in P$ and $v(b)=0$.

Let (X,ℓ_0) be a fixed set of data for the symplectic differential problem in $I\!\!R^n$; recall that there is a subspace $P\subset I\!\!R^n$ and a symmetric bilinear form $S:P\times P\to I\!\!R$ canonically associated to ℓ_0 as in (2.11). We will take into consideration the Hilbert space

[1] This can be done, for instance, in the following way. Consider an arbitrary principal connection on the principal bundle $\mathrm{Sp}(2n,I\!\!R)\to\Lambda$ and let $\hat\Psi_k$ be the parallel lifting of ℓ_k with $\hat\Psi_k(a)=\mathrm{Id}$; then, Ψ_k tends in the C^1-topology to the parallel lifting $\hat\Psi$ of ℓ starting at the identity. Define $\Psi_k=\hat\Psi_k\hat\Psi^{-1}\Psi$.

$\mathcal{H} = H^1_P([a,b];\mathbb{R}^n)$; let I be the following bounded symmetric bilinear form on $\mathcal{H}$:

$$
\begin{aligned}
I(v,w) &= \int_a^b \Big[B(\alpha_v,\alpha_w) + C(v,w)\Big]\,\mathrm{d}t - S(v(a),w(a)) = \\
&= \int_a^b \Big[B^{-1}(v' - Av, w' - Aw) + C(v,w)\Big]\,\mathrm{d}t - S(v(a),w(a)).
\end{aligned}
\tag{2.29}
$$

We will call I the *index form* associated to the pair (X,ℓ_0). For instance, if the symplectic system comes from the Morse–Sturm equation (2.17), then I becomes:

$$
I(v,w) = \int_a^b \Big[g(v',w') + g(Rv,w)\Big]\,\mathrm{d}t - S(v(a),w(a));
\tag{2.30}
$$

this is the classical index form for Morse–Sturm systems.

Integration by parts in (2.29) and the Fundamental Lemma of Calculus of Variations show that:

$$
\mathrm{Ker}(I) = \big\{v \in \mathbb{V} : v(b) = 0\big\},
\tag{2.31}
$$

where $\mathbb{V}$ is defined in (2.13).

Remark 2.5.1. If S is negative semidefinite, B is positive definite and C is positive semidefinite, then I is positive definite in $\mathcal{H}$. By considering restrictions to the interval $[a,t]$, it follows easily from (2.31) that the pair (X,ℓ_0) does not have focal instants in $]\,a,b]$. Analogous results hold if S is positive semidefinite, B is negative definite and C is negative semidefinite.

We will see in Subsection 2.7 that if B is positive definite, then the index $n_-(I)$ in $\mathcal{H}$ is finite, and this number is related to the focal instants of the symplectic differential problem. However, for a non positive B, the index of I in $\mathcal{H}$ is infinite:

Proposition 2.5.2. *The index of I in $\mathcal{H}$ is infinite, unless B is positive definite.*

Proof. First, if v is an X-solution and $f \in H^1_0([a,b];\mathbb{R})$, then, using (2.9), we compute:

$$
I(fv,fv) = \int_a^b (f')^2 B^{-1}(v,v)\,\mathrm{d}t.
\tag{2.32}
$$

Let $t_0 \in\,]a,b[$ and let v be any X-solution such that $B(t_0)^{-1}(v(t_0),v(t_0)) < 0$, so that $B^{-1}(v,v) < 0$ on $[t_0-\varepsilon, t_0+\varepsilon]$ for some $\varepsilon > 0$. It follows from (2.32) that I is negative definite on the infinite dimensional space of maps fv, where f is in $H^1_0([a,b];\mathbb{R})$ and has support in $[t_0-\varepsilon, t_0+\varepsilon]$. □

One of the main goals of the paper is to determine a subspace of $\mathcal{H}$ on which the index of I is finite and that carries the relevant information about the focal instants of the symplectic problem.

2.6. **A subspace where the index is finite.** We want to consider the case that the bilinear form B is not necessarily positive definite; the idea is to *factor out* an infinite dimensional subspace of $\mathcal{H}$ on which I is negative definite, and such that the kernel of I remains unchanged.

Let's assume that the bilinear form B has index k in $\mathbb{R}^{n*}$:

$$
n_-(B) = n_-(B^{-1}) = k.
\tag{2.33}
$$

Let $\mathcal{D} = \{\mathcal{D}_t\}_{t\in[a,b]}$ be a distribution of subspaces $\mathcal{D}_t \subset \mathbb{R}^n$, with $\dim(\mathcal{D}_t) = k$; assume that $B(t)^{-1}$ is negative definite on $\mathcal{D}_t$ for each t, and that $\mathcal{D}_t$ has a C^2-dependence on t,

in the sense that there exists a basis $Y_1(t),\ldots,Y_k(t)$ of $\mathcal{D}_t$, where each Y_i is of class C^2. We introduce the following closed subspace of $\mathcal{H}$:

$$\begin{aligned}\mathcal{K}=\big\{v\in\mathcal{H}:\alpha_v(Y_i)\in H^1([a,b];I\!R)\text{ and}\\ \alpha_v(Y_i)'=B(\alpha_v,\alpha_{Y_i})+C(v,Y_i),\ \forall\, i=1,\ldots,k\big\}.\end{aligned}\tag{2.34}$$

It is not hard to see that $\mathcal{K}$ depends *only* on the distribution of spaces $\mathcal{D}$, and not on the choice of the frame $Y_1,\ldots,Y_k$. Roughly speaking, the space $\mathcal{K}$ is to be interpreted as the space of those $v\in\mathcal{H}$ that are X*-solutions along* $\mathcal{D}$; namely, if $v\in\mathcal{H}$ is of class C^2, then an easy computation shows that $v\in\mathcal{K}$ if and only if:

$$\alpha_v'(Y_i)=\big(Cv-A^*\alpha_v\big)(Y_i),\quad i=1,\ldots,k.$$

In particular, by (2.31) we have:

$$\mathrm{Ker}(I)\subset\mathcal{K}.\tag{2.35}$$

We associate to the choice of $Y_1,\ldots,Y_k$ the following symplectic differential system in $I\!R^k$:

$$\left\{\begin{array}{l}f'=-\mathcal{B}^{-1}\mathcal{C}_{\mathrm{a}}f+\mathcal{B}^{-1}\varphi;\\ \varphi'=(\mathcal{I}-\mathcal{C}_{\mathrm{s}}'+\mathcal{C}_{\mathrm{a}}\mathcal{B}^{-1}\mathcal{C}_{\mathrm{a}})f-\mathcal{C}_{\mathrm{a}}\mathcal{B}^{-1}\varphi,\end{array}\right.\qquad t\in[a,b],\ f(t)\in I\!R^k,\ \varphi(t)\in I\!R^{k^*},\tag{2.36}$$

where $\mathcal{B},\mathcal{I}$ are bilinear forms in $I\!R^k$, and $\mathcal{C},\mathcal{C}_{\mathrm{a}},\mathcal{C}_{\mathrm{s}}$ are linear maps from $I\!R^k$ to $I\!R^{k^*}$, whose matrices in the canonical basis are given by:

$$\begin{aligned}&\mathcal{B}_{ij}=B^{-1}(Y_i,Y_j),\quad \mathcal{C}_{ij}=\alpha_{Y_j}(Y_i),\quad \mathcal{C}_{\mathrm{s}}=\frac{1}{2}(\mathcal{C}+\mathcal{C}^*),\\ &\mathcal{C}_{\mathrm{a}}=\frac{1}{2}(\mathcal{C}-\mathcal{C}^*),\quad \mathcal{I}_{ij}=B(\alpha_{Y_i},\alpha_{Y_j})+C(Y_i,Y_j).\end{aligned}\tag{2.37}$$

We call (2.36) the *reduced symplectic system* associated to (2.4) with respect to $Y_1,\ldots,Y_k$.

We will make the following Assumption for Subsections 2.6, 2.7 and 2.8:

Assumption 2.6.1. *We assume that the symplectic differential system* (2.36) *with initial condition $f(a)=0$ has no focal instants in $]a,b]$, i.e., if f is a non zero solution of* (2.36) *vanishing at $t=a$, then f does not vanish in $]a,b]$.*

Observe that the coefficient $\mathcal{B}^{-1}$ in (2.36) is negative definite, and in Remark 2.5.1 we gave a criterion for the absence of focal instants: if the symmetric bilinear form $\mathcal{I}-\mathcal{C}_{\mathrm{s}}'+\mathcal{C}_{\mathrm{a}}\mathcal{B}^{-1}\mathcal{C}_{\mathrm{a}}$ is negative semidefinite, then Assumption 2.6.1 is satisfied (see also Remark 2.10.5 ahead).

Remark 2.6.2. Although the definition of the system (2.36) depends on the choice of a basis $\{Y_1,\ldots,Y_k\}$ of $\mathcal{D}$, the reduced symplectic systems associated to different choices of frames of $\mathcal{D}$ are *isomorphic*, in a sense that will be clarified in Subsection 2.10 (see Proposition 2.10.4). In particular, the validity of Assumption 2.6.1 does not depend on the choice of $Y_1,\ldots,Y_k$.

The motivation for the definition of the reduced symplectic system is given by Lemma 2.6.8 below.

The reader can keep in mind the following examples where Assumption 2.6.1 holds.

Example 2.6.3. If $k=0$, then there is no need of any assumption. If $k\ge1$ and there exist pointwise linearly independent X-solutions $Y_1,\ldots,Y_k$ such that:

1. B^{-1} is negative definite on the span of $Y_1,\ldots,Y_k$ for all t;
2. the matrix $\alpha_{Y_i}(Y_j)$ is symmetric for all t,

then the right hand side of the second equation of (2.36) vanishes identically, and therefore Assumption 2.6.1 is satisfied. In this case, it is not hard to see that the space $\mathcal{K}$ defined in (2.34) is given by:

$$\mathcal{K} = \big\{ v \in \mathcal{H} : \alpha_v(Y_i) - \alpha_{Y_i}(v) \text{ is constant } \forall\, i = 1, \ldots, k \big\}.$$

For instance, if $k = 1$ then, in order to satisfy Assumption 2.6.1, one only needs to determine an X-solution Y such that $B^{-1}(Y, Y)$ is negative.

Example 2.6.4. If the original symplectic system is a Morse–Sturm system of the form (2.17), then an example where Assumption 2.6.1 holds is obtained when it is possible to find a constant k-dimensional subspace $\mathcal{D}$ on which $B(t)^{-1}$ is negative definite and $C(t)$ is negative semi-definite for all t. In this case, we can take constant vector fields $Y_1, \ldots, Y_k$ as a basis for $\mathcal{D}$.

We now discuss some properties of the space $\mathcal{K}$.

Remark 2.6.5. Let E be a closed subspace of $H^1([a, b]; I\!R^n)$, suppose that $B \in \mathrm{B}(E, I\!R)$ is a bilinear form. Let E_0 denote the normed vector space obtained by considering the C^0-topology on E and assume that B is continuous in $E_0 \times E$. It follows that the linear operator $T \in \mathcal{L}(E)$ which represents B is *compact*, since the inclusion of H^1 in C^0 is compact. The same conclusion holds if we assume that B is continuous in $E \times E_0$, since the adjoint of a compact operator in a Hilbert space is also compact.

We are now ready for the following:

Lemma 2.6.6. *The restriction of I to $\mathcal{K}$ is represented by a compact perturbation of a positive isomorphism.*

Proof. Consider the bilinear form $I_0(v, w) = \int_a^b B^{-1}(v', w')\ \mathrm{d}t$; recalling (2.29), from Remark 2.6.5 it follows easily that $I - I_0$ is represented by a compact operator on $\mathcal{K}$. It remains to prove that I_0 is represented by a compact perturbation of a positive isomorphism of $\mathcal{K}$. For each t, we define a positive definite inner product r_t on $I\!R^n$ by setting

$$r_t(x, y) = B(t)^{-1}(x, y) - 2\, B(t)^{-1}(\pi(t)x, \pi(t)y),$$

where $\pi(t) : I\!R^n \to \mathcal{D}(t)$ is the orthogonal projection with respect to $B(t)^{-1}$. We have:

$$B^{-1}(v, w) = r(v, w) - 2\, r(\pi(v), \pi(w)), \tag{2.38}$$

and it follows:

$$I_0(v, w) = \int_a^b \Big[r(v', w') - 2\, r(\pi(v'), \pi(w')) \Big]\ \mathrm{d}t.$$

Clearly, the integral of the first term above gives a Hilbert space inner product in $\mathcal{K}$, and it is therefore represented by a positive isomorphism of $\mathcal{K}$. Using again Remark 2.6.5, to conclude the proof it suffices to show that the linear map $\mathcal{K} \ni v \mapsto \pi(v') \in L^2([a, b]; I\!R^n)$ is continuous in the C^0-topology of $\mathcal{K}$. To this aim, it clearly suffices to show that the maps $v \mapsto \alpha_v(Y_i)$ are C^0-continuous for all i. This follows from the fact that, for $v \in \mathcal{K}$ the quantity $c_i(v)$ defined by:

$$c_i(v) = \alpha_v(Y_i) - \int_a^t \Big[B(\alpha_v, \alpha_{Y_i}) + C(v, Y_i) \Big]\ \mathrm{d}s \tag{2.39}$$

is constant (see formula (2.34)). Using (2.8) and integration by parts, we see that the integral in (2.39) is continuous in v with respect to the C^0-topology. Integrating (2.39) on $[a, b]$, we see that the functional $\mathcal{K} \ni v \mapsto c_i(v)$ is C^0-continuous, which concludes the proof. $\square$

Corollary 2.6.7. *I has finite index on $\mathcal{K}$.* □

We now describe $\mathcal{K}$ as the kernel of a bounded linear map $F:\mathcal{H}\to L^2([a,b];\mathbb{R}^{k^*})/\mathfrak{C}$, where $\mathfrak{C}$ is the subspace of constant maps. Let $F=(F^1,\ldots,F^k)$ be defined by:

$$F^i(v)(t)=\alpha_v(t)(Y_i(t))-\int_a^t\Big[B(\alpha_v,\alpha_{Y_i})+C(v,Y_i)\Big]\,\mathrm{d}s+\ \text{constant}. \tag{2.40}$$

From (2.34) it follows that $\mathcal{K}=\mathrm{Ker}(F)$; moreover, F is clearly continuous in $\mathcal{H}$, and therefore $\mathcal{K}$ is closed. Next, we introduce a closed subspace $\mathcal{S}$, that will be shown to be a closed complement of $\mathcal{K}$ in $\mathcal{H}$:

$$\mathcal{S}=\big\{v\in H_0^1([a,b];\mathbb{R}^n):\ v(t)\in\mathcal{D}(t),\ \forall\,t\in[a,b]\big\}. \tag{2.41}$$

Lemma 2.6.8. $\mathcal{K}\cap\mathcal{S}=\{0\}$.

Proof. If $v=\sum_i f_iY_i\in\mathcal{S}$, then $v\in\mathcal{K}$ iff $f=(f_1,\ldots,f_k)$ is a solution of the reduced symplectic system (2.36). □

Lemma 2.6.9. *The restriction of F to $\mathcal{S}$ is onto $L^2([a,b];\mathbb{R}^{k^*})/\mathfrak{C}$.*

Proof. We identify $\mathcal{S}$ with $H_0^1([a,b];\mathbb{R}^k)$ by $v=\sum_i f_iY_i\mapsto f=(f_1,\ldots,f_k)$. For $v\in\mathcal{S}$, we can rewrite $F(v)$ with the help of (2.37) as:

$$F(v)(t)=\mathcal{B}(t)f'(t)+\mathcal{C}(t)f(t)-\int_a^t\Big[\mathcal{C}^*f'+\mathcal{I}f\Big]\,\mathrm{d}s+\text{constant}. \tag{2.42}$$

Define $F_0(f)=\mathcal{B}f'$; then, F_0 is a linear isomorphism between the spaces $H_0^1([a,b];\mathbb{R}^k)$ and $L^2([a,b];\mathbb{R}^{k^*})/\mathfrak{C}$. Using integration by parts, the formula (2.42) shows that the linear operator $F-F_0$ is continuous in the C^0-topology, and it is therefore compact. Then, F is a Fredholm operator of index zero on $\mathcal{S}$, and the injectivity of F on $\mathcal{S}$ (Lemma 2.6.8) implies that F is onto. □

Corollary 2.6.10. $\mathcal{H}=\mathcal{K}\oplus\mathcal{S}$.

Proof. It follows from Lemma 2.6.8, Lemma 2.6.9 and the fact that $\mathcal{K}=\mathrm{Ker}(F)$. □

Lemma 2.6.11. *$\mathcal{K}$ and $\mathcal{S}$ are I-orthogonal, i.e., $I(v,w)=0$ for all $v\in\mathcal{S}$ and $w\in\mathcal{K}$.*

Proof. It is an easy computation. □

Corollary 2.6.12. $\mathrm{Ker}(I|_{\mathcal{K}})=\mathrm{Ker(I)}$.

Proof. It follows easily from (2.35), Corollary 2.6.10 and Lemma 2.6.11. □

2.7. **The index theorem.** We have proven the finiteness of the index of I in $\mathcal{K}$; in this section we determine the value of this index in terms of the Maslov index of the pair (X,ℓ_0). To this aim, we study the *evolution* of the index function $i(t)$, $t\in]a,b]$, defined as the index of the form I_t on the space $\mathcal{K}_t\subset\mathcal{H}_t$, where $\mathcal{H}_t=H_P^1([a,t];\mathbb{R}^n)$ and I_t and $\mathcal{K}_t$ are defined as in formulas (2.29) and (2.34) by replacing b with t and $\mathcal{H}$ with $\mathcal{H}_t$. Similarly, one can define in an obvious way the *objects* F_t and $\mathcal{S}_t$ as in (2.40) and (2.41); clearly, all the results proven in Subsection 2.6 remain valid when the symplectic differential system is restricted to the interval $[a,t]$. Observe that for this reason in Assumption 2.6.1 we have required that the reduced symplectic system (2.36) had no focal instants in the entire interval $]a,b]$, and not just at the final instant $t=b$.

We will use the isomorphisms $\Phi_t:\mathcal{H}\to\mathcal{H}_t$ defined by: $\Phi_t(\hat v)=v$, where

$$v(s)=\hat v(u_s),\quad u_s=a+\frac{b-a}{t-a}(s-a),\quad \forall\, s\in[a,t]; \tag{2.43}$$

and we get a family $\{\hat{\mathcal{K}}_t\}_{t\in\,]a,b]}$ of closed subspaces of $\mathcal{H}$, a curve $\hat I:\,]a,b]\to \mathrm{B}_{\mathrm{sym}}(\mathcal{H},I\!\!R)$ of symmetric bilinear forms and a curve $\hat F:\,]a,b]\to\mathcal{L}(\mathcal{H},L^2([a,b];I\!\!R^{k^*})/\mathfrak{C})$ of maps, defined by:

$$\hat{\mathcal{K}}_t=\Phi_t^{-1}(\mathcal{K}_t),\quad \hat I_t=I(\Phi_t\,\cdot\,,\Phi_t\,\cdot\,),\quad \hat F_t=\Phi_t^{-1}\circ F_t\circ\Phi_t.$$

We are also denoting by Φ_t the isomorphism from $L^2([a,b];I\!\!R^{k^*})/\mathfrak{C}$ to $L^2([a,t];I\!\!R^{k^*})/\mathfrak{C}$ defined by formula (2.43). An explicit formula for $\hat I_t$ is given by:

(2.44)
$$\begin{aligned}\hat I_t(\hat v,\hat w)=&\int_a^t B(s)^{-1}\left(\frac{b-a}{t-a}\hat v'(u_s)-A(s)(\hat v(u_s)),\frac{b-a}{t-a}\hat w'(u_s)-A(s)(\hat w(u_s))\right)\mathrm{d}s\\&+\int_a^t C(s)(\hat v(u_s),\hat w(u_s))\,\mathrm{d}s-S(\hat v(a),\hat w(a)).\end{aligned}$$

For $t\in\,]a,b]$, we set $\mathcal{J}_t=(t-a)\hat I_t$:

(2.45)
$$\begin{aligned}\mathcal{J}_t(\hat v,\hat w)=&\int_a^b B(s_u)^{-1}((b-a)\hat v'(u)-(t-a)A(s_u)\hat v(u),\hat w'(u))\,\mathrm{d}u\\&-\frac{1}{b-a}\int_a^b B(s_u)^{-1}((b-a)\hat v'(u)-(t-a)A(s_u)\hat v(u),(t-a)A(s_u)\hat w(u))\,\mathrm{d}u\\&+\frac{1}{b-a}\int_a^b(t-a)^2C(s_u)(\hat v(u),\hat w(u))\,\mathrm{d}u-(t-a)S(\hat v(a),\hat w(a)),\end{aligned}$$

where $s_u=a+\frac{t-a}{b-a}(u-a)$. Setting $t=a$, formula (2.45) defines $\mathcal{J}_a$ as:

$$\mathcal{J}_a(\hat v,\hat w)=(b-a)\int_a^b B(a)^{-1}(\hat v'(u),\hat w'(u))\,\mathrm{d}u. \tag{2.46}$$

Obviously, for $t\in\,]a,b]$, we have

$$i(t)=n_-(I_t|_{\mathcal{K}_t})=n_-(\hat I_t|_{\hat{\mathcal{K}}_t})=n_-(\mathcal{J}_t|_{\hat{\mathcal{K}}_t}). \tag{2.47}$$

In order to study the evolution of the function i, we will make full use of some results presented in reference [11] concerning the jumps of the index of a C^1-curve of bounded symmetric bilinear forms restricted to a C^1-family of closed subspaces of a fixed Hilbert space. For the reader's convenience, we recall some definitions and facts presented in [11].

Definition 2.7.1. Let H be a Hilbert space, $I\subset I\!\!R$ an interval and $\{D_t\}_{t\in I}$ be a family of closed subspaces of H. We say that $\{D_t\}_{t\in I}$ is a C^1-*family* of subspaces if for all $t_0\in I$ there exists a C^1-curve $\beta:\,]t_0-\varepsilon,t_0+\varepsilon[\,\cap I\to\mathcal{L}(H)$ and a closed subspace $\overline{D}\subset H$ such that $\beta(t)$ is an isomorphism and $\beta(t)(D_t)=\overline{D}$ for all t.

We have a method for producing C^1-families of closed subspaces:

Proposition 2.7.2. *Let $I\subset I\!\!R$ be an interval, $H,\tilde H$ be Hilbert spaces and $\mathcal{F}:I\to\mathcal{L}(H,\tilde H)$ be a C^1-map such that each $\mathcal{F}(t)$ is surjective. Then, the family $D_t=\mathrm{Ker}(\mathcal{F}(t))$ is a C^1-family of closed subspaces of H.*

Proof. See [11, Lemma 2.9]. □

Here comes a method for computing the jumps of the index of a smooth curve $M(t)$ of symmetric bilinear forms represented by a compact perturbation of positive isomorphisms.

Proposition 2.7.3. *Let H be a real Hilbert space with inner product $\langle\cdot,\cdot\rangle$, and let $M : [t_0, t_0+r] \to \mathrm{B}_{\mathrm{sym}}(H, I\!R)$, $r > 0$, be a map of class C^1. Let $\{D_t\}_{t\in[t_0,t_0+r]}$ be a C^1-family of closed subspaces of H, and denote by $\overline{M}(t)$ the restriction of $M(t)$ to $D_t \times D_t$. Assume that the following three hypotheses are satisfied:*

1. *$\overline{M}(t_0)$ is represented by a compact perturbation of a positive isomorphism of D_{t_0};*
2. *the restriction $\widetilde{M}$ of the derivative $M'(t_0)$ to $\mathrm{Ker}(\overline{M}(t_0)) \times \mathrm{Ker}(\overline{M}(t_0))$ is non degenerate;*
3. *$\mathrm{Ker}(\overline{M}(t_0)) \subseteq \mathrm{Ker}(M(t_0))$.*

Then, for $t > t_0$ sufficiently close to t_0, $\overline{M}(t)$ is non degenerate, and we have:

$$n_-(\overline{M}(t)) = n_-(\overline{M}(t_0)) + n_-(\widetilde{M}), \tag{2.48}$$

all the terms of the above equality being finite natural numbers.

Proof. See [11, Proposition 2.5]. □

We have the following two immediate corollaries of Proposition 2.7.3.

Corollary 2.7.4. *Under the hypotheses of Proposition 2.7.3, if M and D_t are defined and of class C^1 in a neighborhood of t_0, then, for $\varepsilon > 0$ sufficiently small, we have:*

$$n_-(\overline{M}(t_0+\varepsilon)) - n_-(\overline{M}(t_0-\varepsilon)) = -\mathit{sgn}(\widetilde{M}). \tag{2.49}$$

Proof. Apply twice Proposition 2.7.3, once to M and again to a *backwards* reparameterization of M. □

Corollary 2.7.5. *Under the hypotheses of Proposition 2.7.3, if $\overline{M}(t_0)$ is nondegenerate, then $n_-(\overline{M}(t))$ is constant for t near t_0.* □

We now show how our setup fits into the assumptions of Proposition 2.7.3.

For $t \in\,]a,b]$, we set $\mathcal{F}_t = (t-a)\hat{F}_t$; explicitly, for $i = 1,\dots,k$, $\hat{v} \in \mathcal{H}$ and $u \in [a,b]$, we have:

$$\begin{aligned}\mathcal{F}_t^i(\hat{v})(u) = {} & B(s_u)^{-1}((b-a)\hat{v}(u) - (t-a)A(s_u)\hat{v}(u), Y_i(s_u)) \\ & - \frac{s_u - a}{u-a}\int_a^u \alpha_{Y_i}(r_x)((b-a)\hat{v}'(x) - (t-a)A(r_x)\hat{v}(x))\,\mathrm{d}x \\ & - \frac{s_u-a}{u-a}\int_a^u (t-a)C(r_x)(\hat{v}(x), Y_i(r_x))\,\mathrm{d}x + \text{constant},\end{aligned} \tag{2.50}$$

where $r_x = a + \frac{s_u-a}{u-a}(x-a)$. Setting $t = a$, formula (2.50) defines $\mathcal{F}_a$ as:

$$\mathcal{F}_a^i(\hat{v})(u) = (b-a)B(a)^{-1}(\hat{v}'(u), Y_i(a)) + \text{constant}. \tag{2.51}$$

Obviously, $\mathrm{Ker}(\mathcal{F}_t) = \hat{\mathcal{K}}_t$ for $t \in\,]a,b]$; we set $\hat{\mathcal{K}}_a = \mathrm{Ker}(\mathcal{F}_a)$, namely:

$$\hat{\mathcal{K}}_a = \big\{\hat{v} \in \mathcal{H} : B(a)^{-1}(\hat{v}'(u), Y_i(a)) = \text{constant},\ i = 1,\dots,k\big\}. \tag{2.52}$$

Proposition 2.7.6. *Assume that C is a map of class C^1. Then, $\hat{I}$ is a map of class C^1 in $]a,b]$, $\mathcal{J}$ is of class C^1 in $[a,b]$ and $\{\hat{\mathcal{K}}_t\}_{t\in[a,b]}$ is a C^1-family of closed subspaces of $\mathcal{H}$.*

Proof. By standard regularity arguments (see [11, Lemma 2.3, Proposition 3.3 and Lemma 4.3]), formula (2.45) shows that $\mathcal{J}$ is C^1 in $[a,b]$, which obviously implies that $\hat{I}$ is C^1 in $]a,b]$. Similarly, formula (2.50) shows that $\mathcal{F}$ is of class C^1 on $[a,b]$; from Lemma 2.6.9 we deduce that $\mathcal{F}_t$ is surjective for $t\in\,]a,b]$. The surjectivity of $\mathcal{F}_a$ is immediately established from its definition (2.51). The regularity of the family $\{\hat{\mathcal{K}}_t\}_{t\in[a,b]}$ follows then from Proposition 2.7.2. □

The last ingredient missing for applying Proposition 2.7.3 to the curve $\hat{I}$ is the computation of the derivative of $\hat{I}$ on its kernel. This is done in the following:

Lemma 2.7.7. *For each $t\in\,]a,b]$, the map $\sigma_t:\mathrm{Ker}(I_t)\to\mathbb{V}[t]^o$ given by $\sigma_t(v)=\alpha_v(t)$ is an isomorphism. Moreover, if C is of class C^1, the isomorphism $\sigma_t\circ\Phi_t$ carries the restriction of the derivative $\hat{I}_t'$ to $\mathrm{Ker}(\hat{I}_t)$ into the restriction of $-B(t)$ to $\mathbb{V}[t]^o$.*

Proof. The fact that σ_t is an isomorphism follows easily from formulas (2.15) and (2.31).

Let $\hat{v},\hat{w}\in\mathrm{Ker}(\hat{I}_t)$; set $v=\Phi_t(\hat{v})$ and $w=\Phi_t(\hat{w})$, so that $v,w\in\mathrm{Ker}(I_t)$. From (2.44), considering that $v(t)=w(t)=0$ we compute:

(2.53)
$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}\,\hat{I}_t(\hat{v},\hat{w})=\\
&\quad B(t)^{-1}(v'(t),w'(t))-\frac{1}{t-a}\int_a^t\Big[2B^{-1}(v',w')-B^{-1}(v',Aw)-B^{-1}(Av,w')\Big]\,\mathrm{d}s\\
&-\frac{1}{t-a}\int_a^t(s-a)\Big[B^{-1}(v'',w')+B^{-1}(v',w'')-B^{-1}(v'',Aw)-B^{-1}(v',Aw')\Big]\mathrm{d}s\\
&\quad-\frac{1}{t-a}\int_a^t(s-a)\Big[-B^{-1}(Av',w')-B^{-1}(Av,w'')+B^{-1}(Av',Aw)\Big]\,\mathrm{d}s\\
&\quad-\frac{1}{t-a}\int_a^t(s-a)\Big[B^{-1}(Av,Aw')+C(v',w)+C(v,w')\Big]\,\mathrm{d}s,
\end{aligned}$$

where all the functions inside the integrals above are meant to be evaluated at s. We use the equation (2.19) satisfied by v and w to eliminate the two terms containing the bilinear form C in (2.53), and we obtain:

(2.54)
$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\hat{I}_t(\hat{v},\hat{w})&=B(t)^{-1}(v'(t),w'(t))\\
&\quad-\frac{1}{t-a}\int_a^t\frac{\mathrm{d}}{\mathrm{d}s}\Big[(s-a)\big(B^{-1}(v',w'-Aw)+B^{-1}(v'-Av,w')\big)\Big]\,\mathrm{d}s=\\
&=-B(t)^{-1}(v'(t),w'(t))=-B(t)(\alpha_v(t),\alpha_w(t)).
\end{aligned}$$

This concludes the proof. □

We now consider the case $t=a$.

Lemma 2.7.8. *The restriction of the symmetric bilinear form $\mathcal{J}_a$ to $\hat{\mathcal{K}}_a$ is represented by a compact perturbation of a positive isomorphism. Moreover, it is non degenerate, and its index equals the index of the restriction of $B(a)^{-1}$ to P.*

Proof. The fact that the restriction of $\mathcal{J}_a$ to $\hat{\mathcal{K}}_a$ is represented by a compact perturbation of a positive isomorphism follows by arguments similar to those used in the proof of Lemma 2.6.6.

Write $P = P_+ \oplus P_-$, where $B(a)^{-1}$ is positive definite on P_+ and negative definite on P_-; this is possible because, by Assumption 2.1.2, $B(a)^{-1}$ is nondegenerate on P. We now define the following subspaces of $\hat{\mathcal{K}}_a$:

$$\hat{\mathcal{K}}_+ = \big\{\hat{v} \in \hat{\mathcal{K}}_a : \hat{v}(a) \in P_+\big\},$$
$$\hat{\mathcal{K}}_- = \big\{\hat{v} : \hat{v} \text{ is an affine function, } \hat{v}(a) \in P_-,\ \hat{v}(b) = 0\big\}. \tag{2.55}$$

It is easily seen that $\hat{\mathcal{K}}_a = \hat{\mathcal{K}}_+ \oplus \hat{\mathcal{K}}_-$; we claim that $\mathcal{J}_a$ is positive definite on $\hat{\mathcal{K}}_+$ and negative definite on $\hat{\mathcal{K}}_-$. Namely, for $v_0 \in P_-$, $v_0 \neq 0$, and $\hat{v} \in \hat{\mathcal{K}}_-$ of the form $\hat{v}(u) = v_0(u-a)$, from (2.46) it is easily computed $\mathcal{J}_a(\hat{v},\hat{v}) = (b-a)^2 B(a)^{-1}(v_0,v_0) < 0$. Suppose now $\hat{v} \in \hat{\mathcal{K}}_+$. Using the Cauchy–Schwarz inequality componentwise, it is easily proven the following inequality for any integrable function $z : [a,b] \to I\!R^n$:

$$r_a\left(\int_a^b z, \int_a^b z\right) \le (b-a)\int_a^b r_a(z,z), \tag{2.56}$$

where the equality holds if and only if z is constant almost everywhere. Applying (2.56) to $\hat{v}'$, we get:

$$\begin{aligned}\mathcal{J}_a(\hat{v},\hat{v}) \ge r_a\left(\int_a^b \hat{v}', \int_a^b \hat{v}'\right) - 2r_a(\pi(a)(\hat{v}(a)), \pi(a)(\hat{v}(a))) =\\ = B(a)^{-1}(\hat{v}(a),\hat{v}(a)) \ge 0,\end{aligned} \tag{2.57}$$

where equality between the first and the last term above occurs if and only if $\hat{v}$ is affine and $\hat{v}(a) = 0$, i.e., if and only if $\hat{v} = 0$. This proves the claim; it follows easily that $\mathcal{J}_a$ is nondegenerate on $\hat{\mathcal{K}}_a$ and that its index is equal to $\dim(\hat{\mathcal{K}}_-) = \dim(P_-) = n_-(B(a)^{-1}|_P)$. □

We are ready to prove the index theorem for nondegenerate pairs (X, ℓ_0) (recall Subsection 2.4).

Lemma 2.7.9. *Let (X,ℓ_0) be a nondegenerate pair, with X of class C^1. Assume that $t = b$ is not a focal instant. Then,*

$$n_-(I|_{\mathcal{K}}) = n_-(B(a)^{-1}|_P) + \mathrm{i}_{\mathrm{foc}}(X,\ell_0). \tag{2.58}$$

Proof. By Theorem 2.3.3, the number of focal instants is finite. Using (2.31) and Corollary 2.6.12, we see that $\hat{I}_t$ is nondegenerate on $\hat{\mathcal{K}}_t$ when t is not focal. Applying Corollary 2.7.5, where Lemma 2.6.6 is being used, the function $i(t)$ defined in (2.47) is piecewise constant, more precisely, it is constant on every interval that does not contain focal instants. Using Corollary 2.7.4 and Lemma 2.7.7, the jump of $i(t)$ at a focal instant $t \in \,]a,b[$ is equal to *sgn*(t). Finally, by Corollary 2.7.5 and Lemma 2.7.8, $i(t) = n_-(B(a)^{-1}|_P)$ for t sufficiently close to a. This concludes the proof. □

We can now prove the aimed index theorem, which we state in a complete form for future reference:

Theorem 2.7.10 (Index Theorem)**.** *Let (X,ℓ_0) be a set of data for the symplectic differential problem in $I\!R^n$, with A, B of class C^1, C continuous, $B(t)$ invertible for all t, and let $\{\mathcal{D}_t\}_{t\in[a,b]}$ be a C^2-family of k-dimensional subspaces of $I\!R^n$, where $k = n_-(B(t))$, and $B(t)^{-1}$ is negative definite on $\mathcal{D}_t$ for all t. Suppose that Assumption 2.1.2 and Assumption 2.6.1 are satisfied, and that $t = b$ is not a focal instant. Define $\mathcal{K}$ as in* (2.34)*; then, the index of I on $\mathcal{K}$ is finite, and the following equality holds:*

$$n_-(I|_{\mathcal{K}}) = n_-(B(a)^{-1}|_P) + \mathrm{i}_{\mathrm{maslov}}(X,\ell_0). \tag{2.59}$$

Proof. Using Theorem 2.3.3 and Lemma 2.7.9, the theorem holds if C is a map of class C^1 and if (X, ℓ_0) is nondegenerate. The conclusion follows from Proposition 2.4.1, observing that all the numbers in formula (2.59) are stable by uniformly small perturbations of the data. □

Remark 2.7.11. If C is of class C^1, the result of Theorem 2.7.10 can be extended to the case that $t = b$ is a nondegenerate focal instant. In this case, $t = b$ is an isolated focal instant, and one can define the Maslov index of the pair (X, ℓ_0), in analogy with formula (2.25), as:

$$\mathrm{i}_{\mathrm{maslov}} = \mu_{L_0}(\ell|_{[a+\varepsilon, b-\varepsilon]}), \tag{2.60}$$

where $\varepsilon > 0$ is small enough. In this case, another application of Proposition 2.7.3 around $t = b$ gives us the following equality:

$$n_-(I|_{\mathcal{K}}) = n_-(B(a)^{-1}|_P) - n_-(B(b)^{-1}|_{\mathbb{V}[b]^o}) + \mathrm{i}_{\mathrm{maslov}}(X, \ell_0). \tag{2.61}$$

2.8. **The case of a variable endpoint.** Motivated by the geometric problem of solutions of a Hamiltonian system with both endpoints variable, we will now consider the following setup.

Given a pair (X, ℓ_0) of data for the symplectic differential problem (2.4), we will additionally consider a Lagrangian subspace ℓ_1 of $(R^n \oplus I\!R^{n*}, \omega)$, to which we associate a pair (Q, S_Q) consisting of a subspace Q of $I\!R^n$ and a symmetric bilinear form S_Q on Q as in (2.11). Given the triple (X, ℓ_0, ℓ_1), we will consider a Hilbert space $\mathcal{H}^\#$ and a bounded symmetric bilinear form $I^\#$ on $\mathcal{H}^\#$ as follows:

$$\mathcal{H}^\# = \big\{ v \in H^1([a, b]; I\!R^n) : v(a) \in P,\ v(b) \in Q \big\}; \tag{2.62}$$

and

$$I^\#(v, w) = \int_a^b \Big[B(\alpha_v, \alpha_w) + C(v, w) \Big]\, \mathrm{d}t + S_Q(v(b), w(b)) - S(v(a), w(a)), \tag{2.63}$$

with α_v, α_w given by (2.8). Comparing with (2.29), clearly $I^\#$ coincides with I on $\mathcal{H}$.

We want to extend the index theorem to this situation, and to this aim we define a space $\mathcal{K}^\#$ in analogy with (2.34):

$$\begin{aligned} \mathcal{K}^\# = \big\{ v \in \mathcal{H}^\# : \alpha_v(Y_i) \in H^1([a, b]; I\!R) \text{ and } \\ \alpha_v(Y_i)' = B(\alpha_v, \alpha_{Y_i}) + C(v, Y_i),\ \ \forall\, i = 1, \dots, k \big\}. \end{aligned} \tag{2.64}$$

where $Y_1, \dots, Y_k$ is a frame for $\mathcal{D}$..

Recalling the curve $\ell(t)$ of Lagrangians (2.21) associated to the pair (X, ℓ_0), we consider the pair $(\mathbb{V}_b, S_b)$ consisting of a subspace $\mathbb{V}_b \subset I\!R^n$ and a symmetric bilinear form S_b on $\mathbb{V}_b$ associated to the Lagrangian $\ell(b)$; explicitly, we have:

$$\mathbb{V}_b = \big\{ v(b) : v \in \mathbb{V} \big\}, \tag{2.65}$$

and

$$S_b(v(b), w(b)) = -\alpha_v(b)(w(b)), \quad \forall\, v, w \in \mathbb{V}. \tag{2.66}$$

If Q is contained in $\mathbb{V}_b$, for instance this is always the case if $t = b$ is not (X, ℓ_0)-focal, then we define a symmetric bilinear form $\mathfrak{Q}$ on Q:

$$\mathfrak{Q} = S_Q - S_b|_Q. \tag{2.67}$$

Theorem 2.8.1. *Under the hypotheses of Theorem 2.7.10, the index of $I^\#$ on $\mathcal{K}^\#$ is finite, and we have:*

$$n_-(I^\#|_{\mathcal{K}^\#}) = n_-(B(a)^{-1}|_P) + n_-(\mathfrak{Q}) + \mathrm{i}_{\mathrm{maslov}}(X,\ell_0). \tag{2.68}$$

Proof. We define the following finite dimensional vector space:

$$\mathbb{V}_Q = \big\{v \in \mathbb{V} : v(b) \in Q\big\}. \tag{2.69}$$

Since $t = b$ is not focal, it is easily seen that we have a direct sum decomposition $\mathcal{K}^\# = \mathcal{K} \oplus \mathbb{V}_Q$. Moreover, integration by parts in (2.63) shows that, for $v \in \mathbb{V}_Q$ and $w \in \mathcal{H}^\#$, it is:

$$I^\#(v,w) = \mathfrak{Q}(v(b), w(b)). \tag{2.70}$$

This shows that the spaces $\mathcal{K}$ and $\mathbb{V}_Q$ are $I^\#$-orthogonal, and therefore $n_-(I^\#|_{\mathcal{K}^\#}) = n_-(I|_{\mathcal{K}}) + n_-(I^\#|_{\mathbb{V}_Q})$. The conclusion follows from Theorem 2.7.10 and from the observation that the isomorphism $\mathbb{V}_Q \ni v \mapsto v(b) \in Q$ carries $I^\#$ into $\mathfrak{Q}$. $\square$

2.9. **The opposite symplectic system.** It will be useful in the rest of the paper to introduce an operation of *opposition* in the category of symplectic differential systems, that is described briefly in this subsection.

Let (X,ℓ_0) be a set of data for the symplectic differential problem in $\mathbb{R}^n$; we define a new pair $(X^{\mathrm{op}},\ell_0^{\mathrm{op}})$, called the *opposite symplectic differential problem*, by setting:

$$X^{\mathrm{op}} = \mathcal{O}X\mathcal{O}, \quad \ell_0^{\mathrm{op}} = \mathcal{O}(\ell_0), \tag{2.71}$$

where $\mathcal{O} : \mathbb{R}^n \oplus \mathbb{R}^{n*} \to \mathbb{R}^n \oplus \mathbb{R}^{n*}$ is given by $\mathcal{O}(v,\alpha) = (v,-\alpha)$. Opposite symplectic systems have opposite Maslov indexes and opposite index forms; applying Theorem 2.7.10 to $(X^{\mathrm{op}},\ell_0^{\mathrm{op}})$ we get:

Theorem 2.9.1. *Let (X,ℓ_0) be a set of data for the symplectic differential problem in $\mathbb{R}^n$, with A, B of class C^1, C continuous, $B(t)$ invertible for all t, and let $\{\mathcal{D}_t\}_{t\in[a,b]}$ be a C^2-family of k-dimensional subspaces of $\mathbb{R}^n$, where $k = n_+(B(t))$, and $B(t)^{-1}$ is positive definite on $\mathcal{D}_t$ for all t. Suppose that Assumption 2.1.2 and Assumption 2.6.1 are satisfied, and that $t = b$ is not a focal instant. Define $\mathcal{K}$ as in* (2.34)*; then, the co-index of I on $\mathcal{K}$ is finite, and the following equality holds:*

$$n_+(I|_{\mathcal{K}}) = n_+(B(a)^{-1}|_P) - \mathrm{i}_{\mathrm{maslov}}(X,\ell_0). \tag{2.72}$$

2.10. **Equivalence of Symplectic Differential Systems.** In this subsection we describe the natural isomorphisms in the class of symplectic differential problems and we prove the invariance of the index form by such isomorphisms.

Let L_0 be the Lagrangian subspace $\{0\} \oplus \mathbb{R}^{n*}$ of $(\mathbb{R}^n \oplus \mathbb{R}^{n*},\omega)$; we denote by $\mathrm{Sp}(2n,\mathbb{R};L_0)$ the closed subgroup of $\mathrm{Sp}(2n,\mathbb{R})$ consisting of those symplectomorphisms ϕ_0 such that $\phi_0(L_0) = L_0$. It is easily seen that any such symplectomorphism is given in block matrix by:

$$\phi_0 = \begin{pmatrix} Z & 0 \\ Z^{*-1}W & Z^{*-1} \end{pmatrix}, \tag{2.73}$$

with $Z : \mathbb{R}^n \to \mathbb{R}^n$ an isomorphism and W a symmetric bilinear form in $\mathbb{R}^n$.

We give the following:

Definition 2.10.1. The symplectic differential systems with coefficient matrices X and $\tilde{X}$ are said to be *isomorphic* if there exists a C^1-map $\phi_0 : [a,b] \to \mathrm{Sp}(2n, I\!R; L_0)$ whose upper-left $n \times n$ block is of class C^2 and such that:

$$\tilde{X} = \phi_0' \phi_0^{-1} + \phi_0 X \phi_0^{-1}; \tag{2.74}$$

the pairs (X, ℓ_0) and $(\tilde{X}, \tilde{\ell}_0)$ are called *isomorphic* if in addition $\tilde{\ell}_0 = \phi_0(a)(\ell_0)$.

We call the map ϕ_0 an *isomorphism* between X and $\tilde{X}$, or between the pairs (X, ℓ_0) and $(\tilde{X}, \tilde{\ell}_0)$ in the second case.

Denoting by Ψ and $\tilde{\Psi}$ the fundamental matrices of the symplectic systems with coefficient matrices X and $\tilde{X}$ respectively, we have:

$$\tilde{\Psi}(t) = \phi_0(t)\Psi(t)\phi_0(a)^{-1}, \quad \forall\, t \in [a,b]. \tag{2.75}$$

Formula (2.75) is the motivation for Definition 2.10.1: the map ϕ_0 has to be interpreted as a change of variable in the differential system. Similarly, if ℓ and $\tilde{\ell}$ denote the curve of Lagrangians associated to the symplectic problems with data (X, ℓ_0) and $(\tilde{X}, \tilde{\ell}_0)$ respectively, it is easily seen that:

$$\tilde{\ell}(t) = \phi_0(t)(\ell(t)), \quad \forall\, t \in [a,b]. \tag{2.76}$$

Given an isomorphism ϕ_0 between (X, ℓ_0) and $(\tilde{X}, \tilde{\ell}_0)$, we define a C^2-curve Z and a C^1-curve W by formula (2.73). The isomorphy between the pairs can be given in terms of the matrix blocks defining X and $\tilde{X}$ (recall formula (2.3)) as follows:

$$\begin{aligned} \tilde{A} &= ZAZ^{-1} - ZBWZ^{-1} + Z'Z^{-1}, \\ \tilde{B} &= ZBZ^*, \\ \tilde{C} &= Z^{*-1}(WA + C - WBW + A^*W + W')Z^{-1}. \end{aligned} \tag{2.77}$$

It is easily seen that v is a (X, ℓ_0)-solution if and only if Zv is a $(\tilde{X}, \tilde{\ell}_0)$ solution. Also, it follows that isomorphic symplectic differential problems have the same focal instants, and by (2.77) they have the same signature and multiplicity. In particular, isomorphic systems have the same focal index. We also give the formulas for the objects $\tilde{P}$ and $\tilde{S}$ defined by $\tilde{\ell}_0$ as in (2.11):

$$\tilde{P} = Z(a)(P), \quad \tilde{S} = S(Z(a)^{-1}\cdot, Z(a)^{-1}\cdot) - W(a)(Z(a)^{-1}\cdot, Z(a)^{-1}\cdot)|_{\tilde{P}}. \tag{2.78}$$

We now prove:

Proposition 2.10.2. *Let (X, ℓ_0) and $(\tilde{X}, \tilde{\ell}_0)$ be isomorphic pairs such that the final instant is not focal. Then,* $\mathrm{i}_{\mathrm{maslov}}(X, \ell_0) = \mathrm{i}_{\mathrm{maslov}}(\tilde{X}, \tilde{\ell}_0)$.

Proof. Choose $\varepsilon > 0$ small enough so that both ℓ and $\tilde{\ell}$ do not intercept $\Lambda_{\geq 1}(L_0)$ on $]a, a+\varepsilon]$. We prove that $\ell|_{[a+\varepsilon,b]}$ and $\tilde{\ell}|_{[a+\varepsilon,b]}$ define the same homology class in $H_1(\Lambda, \Lambda_0(L_0))$. To this aim, we first observe that the curve ϕ_0 is homotopic to the constant map $\phi_0(b)$ in $\mathrm{Sp}(2n, I\!R; L_0)$. It follows from (2.76) that $\tilde{\ell}$ is homotopic to $\phi_0(b) \circ \ell$ in $(\Lambda, \Lambda_0(L_0))$. It remains to show that the fixed symplectomorphism $\phi_0(b)$ indices the identity map in the relative homology group $H_1(\Lambda, \Lambda_0(L_0))$. This follows easily from the fact that, since $\mathrm{Sp}(2n, I\!R)$ is connected[2], then $\phi_0(b)$ induces the identity map in the absolute homology group $H_1(\Lambda)$. By functoriality of singular homology, it follows that $\phi_0(b)$ also induces the identity map in the relative homology group $H_1(\Lambda, \Lambda_0(L_0))$. □

[2]Observe that the group $\mathrm{Sp}(2n, I\!R; L_0)$ is *not* connected, hence the homotopy argument cannot be done directly at the relative homology level.

Finally, we state the invariance of the index form and of the space $\mathcal{K}$.

Proposition 2.10.3. *Let (X, ℓ_0) and $(\tilde{X}, \tilde{\ell}_0)$ be isomorphic pairs, with associated index forms $I : \mathcal{H} \times \mathcal{H} \to I\!R$ and $\tilde{I} : \tilde{\mathcal{H}} \times \tilde{\mathcal{H}} \to I\!R$ respectively. Then, the operator $\mathcal{H} \ni v \mapsto Zv \in \tilde{\mathcal{H}}$ is a Hilbert space isomorphism that carries I into $\tilde{I}$ and $\mathcal{K}$ onto $\tilde{\mathcal{K}}$.*

Proof. It is an easy computation that uses (2.77) and (2.78). □

Similarly, by a direct computation it is easy to prove the following:

Proposition 2.10.4. *Let X be the coefficient matrix of a symplectic differential system in $I\!R^n$, and let $\{\mathcal{D}_t\}_{t\in[a,b]}$ be a C^2-family of k-dimensional subspaces of $I\!R^n$ such that $B(t)$ is negative definite on $\mathcal{D}_t$ for all t, and $k = n_-(B)$. Then, any two reduced symplectic systems associated to different choices of bases for $\mathcal{D}$ are isomorphic.* □

Remark 2.10.5. In order to ease the verification of Assumption 2.6.1, one could study alternative reduced systems which are isomorphic to (2.36). An example of such alternative is given by the following:

$$\begin{cases} f' = -\mathcal{B}^{-1}\mathcal{C}\, f + \mathcal{B}^{-1}\,\varphi, \\ \varphi' = (\mathcal{I} - \mathcal{C}^*\mathcal{B}^{-1}\mathcal{C})\, f + \mathcal{C}^*\mathcal{B}^{-1}\,\varphi. \end{cases} \tag{2.79}$$

A sufficient condition for a symplectic differential problem not have focal instants is given in Remark 2.5.1. We remark that this condition is *not* preserved by isomorphisms; hence, another sufficient condition that guarantees the validity of Assumption 2.6.1 is that the symmetric bilinear form $\mathcal{I} - \mathcal{C}^*\mathcal{B}^{-1}\mathcal{C}$ be negative semi-definite.

Proposition 2.10.3 and Proposition 2.10.4 show that the index form of any reduced symplectic system defines a unique symmetric bilinear form on the space $\mathcal{S}$ of $\mathcal{D}$-valued vector fields vanishing at the endpoints. A simple computation shows that this reduced index form is the restriction of the original index form:

Proposition 2.10.6. *Let (X, ℓ_0) be a set of data for the symplectic differential problem in $I\!R^n$, $\mathcal{D}$ be a C^2-family of subspaces of $I\!R^n$, and let $Y_1, \ldots, Y_k$ be a basis of $\mathcal{D}$. Denote by I the index form of (X, ℓ_0) and by I_{red} the index form of the reduced symplectic system of (X, ℓ_0) associated to the given basis. Then, for every $f, g \in H^1_0([a, b]; I\!R^k)$, we have:*

$$I_{\mathrm{red}}(f, g) = I(v, w), \tag{2.80}$$

where $v = \sum_i f_i Y_i$ and $w = \sum_i g_i Y_i$. □

Corollary 2.10.7. *Assumption 2.6.1 is equivalent to the condition that the restriction of the index form I to the space $\mathcal{S}$ (defined in* (2.41)*) be negative definite.*

Proof. Apply Theorem 2.9.1 to the reduced symplectic differential system, keeping in mind Proposition 2.10.6. □

2.11. **Every symplectic system is isomorphic to a Morse–Sturm system.** It is a trivial observation that every second order linear differential equation $x'' + f_1 x' + f_2 x = 0$ can be written as a Sturm equation: $(e^{\int f_1} x')' = -f_2 e^{\int f_1} x$. On the other hand, given any Sturm equation $(px')' = rx$, with p a map of class C^2 and, say, $p > 0$, then the change of variable $y = \sqrt{p}x$ transforms the equation into the Sturm equation $y'' = \tilde{r}y$, with $\tilde{r} = \left[\frac{r}{p} + \frac{p''p - \frac{1}{2}(p')^2}{2p^2}\right]$.

In the language of symplectic systems, these facts are expressed by saying that every C^2 unidimensional symplectic system is isomorphic to a Morse–Sturm system having coefficient matrix $\begin{pmatrix} 0 & b \\ c & 0 \end{pmatrix}$ with b constant.

This fact holds for symplectic systems of any dimension:

Proposition 2.11.1. *Every symplectic differential system* (2.4) *in* $\mathbb{R}^n$ *is isomorphic to a Morse–Sturm system, i.e., a symplectic system with coefficient matrix* $\tilde{X} = \begin{pmatrix} 0 & \tilde{B} \\ \tilde{C} & 0 \end{pmatrix}$. *Moreover, if* B *is a map of class* C^2*, then* $\tilde{B}$ *can be taken to be the* constant *map*

$$\begin{pmatrix} -\mathrm{Id}_k & 0 \\ 0 & \mathrm{Id}_{n-k} \end{pmatrix}, \tag{2.81}$$

where $k = n_-(B)$.

Proof. We will use the notation of Subsection 2.10 to exhibit the required isomorphisms. For the first part of the statement, simply consider the symmetric matrix $W = 0$ and let Z be the solution on $[a, b]$ of the initial value problem $Z' = -ZA$ with $Z(a) = \mathrm{Id}$. Then, by formula (2.77), we have $\tilde{A} = 0$.

For the second part of the thesis, observe first that if B is a map of class C^2, then we can assume that B is the constant map (2.81), observing that we can always find a basis $\{e_1(t), \ldots, e_n(t)\}$ in $\mathbb{R}^n$ such that $B(t)$ is in the canonical diagonal form (2.81) and such that each $e_i(t)$ is of class C^2.

Denote by $\mathrm{O}(k, n-k)$ the closed subgroup of $\mathrm{GL}(n, \mathbb{R})$ consisting of those maps preserving the quadratic form with matrix (2.81) and by $\mathrm{o}(k, n-k)$ its Lie algebra. Then, if ϕ_0 is an isomorphism between the symplectic systems X and $\tilde{X}$ such that the corresponding map Z takes values in $\mathrm{O}(k, n-k)$, it follows that $\tilde{B} = ZBZ^* = B$ is also equal to (2.81). Now, if we consider the symmetric matrix $W = \frac{1}{2}(BA + A^*B)$ and if we take Z to be the solution of the initial value problem $Z' = Z(BW - A)$, $Z(a) = \mathrm{Id}$, then $\tilde{A} = 0$, and, since $BW - A \in \mathrm{o}(k, n-k)$, we obtain $Z \in \mathrm{O}(k, n-k)$. This concludes the proof. $\square$

Examples of Morse–Sturm systems in $\mathbb{R}^n$ whose coefficient matrix has constant upper-right $n \times n$ block arise from Jacobi equations along semi–Riemannian geodesics by a parallel trivialization of the tangent bundle along the geodesic (see Subsection 3.5). Conversely, every Morse–Sturm system of this type corresponds to the Jacobi equation along a semi–Riemannian geodesic:

Proposition 2.11.2. *Every smooth Morse–Sturm system in* $\mathbb{R}^n$ *whose coefficient matrix has upper-right* $n \times n$ *block constant can be obtained as the Jacobi equation along a geodesic* γ *in a semi-Riemannian manifold* (M, g) *by a parallel trivialization of the tangent bundle* TM *along* γ.

Proof. See [13, Section 3] and also [19, Proposition 2.3.1]. $\square$

3. Applications to Hamiltonian Systems

In this section we will consider the following setup. Let $(\mathcal{M}, \omega)$ be a symplectic manifold, i.e., $\mathcal{M}$ is a smooth manifold and ω is a smooth closed skew-symmetric nondegenerate two-form on $\mathcal{M}$; we set $\dim(\mathcal{M}) = 2n$. Let $H : U \to \mathbb{R}$ be a smooth function defined in an open set $U \subseteq \mathbb{R} \times \mathcal{M}$; we will call such function a *Hamiltonian* in $(\mathcal{M}, \omega)$. For each $t \in \mathbb{R}$, we denote by H_t the map $m \mapsto H(t, m)$ defined in the open set $U_t \subseteq \mathcal{M}$ consisting of those $m \in \mathcal{M}$ such that $(t, m) \in U$. We denote by $\vec{H}$ the smooth time-dependent vector field in $\mathcal{M}$ defined by $\mathrm{d}H_t(m) = \omega(\vec{H}(t, m), \cdot)$ for all $(t, m) \in U$; let F denote the maximal flow of the vector field $\vec{H}$ defined on an open set of $\mathbb{R} \times \mathbb{R} \times \mathcal{M}$ taking values in $\mathcal{M}$, i.e., for each $m \in \mathcal{M}$ and $t_0 \in \mathbb{R}$, the curve $t \mapsto F(t, t_0, m)$ is a maximal integral

curve of $\vec{H}$ and $F(t_0, t_0, m) = m$. This means that $F(\cdot, t_0, m)$ is a maximal solution of the equation:

$$\frac{\mathrm{d}}{\mathrm{d}t} F(t, t_0, m) = \vec{H}(t, F(t, t_0, m)), \quad F(t_0, t_0, m) = m.$$

Recall that F is a smooth map; we also write F_{t,t_0} for the map $m \mapsto F(t, t_0, m)$; observe that F_{t,t_0} is a diffeomorphism between open subsets of $\mathcal{M}$.

We recall that a *symplectic chart* in $\mathcal{M}$ is a local chart (q, p) taking values in $\mathbb{R}^n \oplus \mathbb{R}^{n*}$ whose differential at each point is a symplectomorphism from the tangent space of $\mathcal{M}$ to $\mathbb{R}^n \oplus \mathbb{R}^{n*}$ endowed with the canonical symplectic structure. We write $q = (q_1, \ldots, q_n)$ and $p = (p_1, \ldots, p_n)$; we denote by $\{\frac{\partial}{\partial q_i}, \frac{\partial}{\partial p_j}\}$, $i, j = 1, \ldots, n$ the corresponding local referential of $T\mathcal{M}$, and by $\{\mathrm{d}q_i, \mathrm{d}p_j\}$ the local referential of $T\mathcal{M}^*$. By Darboux's Theorem, there always exists an atlas of symplectic charts.

In a given symplectic chart (q, p), we have:

$$\omega = \sum_{i=1}^{n} \mathrm{d}q_i \wedge \mathrm{d}p_i, \quad \vec{H} = \sum_{i=1}^{n} \left(\frac{\partial H}{\partial p_i} \frac{\partial}{\partial q_i} - \frac{\partial H}{\partial q_i} \frac{\partial}{\partial p_i} \right).$$

Let $\mathcal{P}$ be a *Lagrangian submanifold* of $\mathcal{M}$, i.e., $T_m\mathcal{P}$ is a Lagrangian subspace of $T_m\mathcal{M}$ for every $m \in \mathcal{P}$. We fix an integral curve $\Gamma : [a, b] \to \mathcal{M}$ of $\vec{H}$, so that $\Gamma(t) = F(t, a, \Gamma(a))$ for all $t \in [a, b]$. We also say that Γ is a *solution of the Hamilton equations*, i.e., in a symplectic chart $\Gamma(t) = (q(t), p(t))$:

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\partial H}{\partial p}, \quad \frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial q}. \tag{3.1}$$

We assume that Γ starts at $\mathcal{P}$, that is $\Gamma(a) \in \mathcal{P}$. Finally, we will consider a fixed smooth distribution $\mathfrak{L}$ in $\mathcal{M}$ such that $\mathfrak{L}_m$ is a Lagrangian subspace of $T_m\mathcal{M}$ for all $m \in \mathcal{M}$.

The basic example to keep in mind for the above setup is the case where $\mathcal{M}$ is the cotangent bundle TM^* of some smooth manifold M endowed with the canonical symplectic structure, $\mathcal{P}$ is the annihilator TP^o of some smooth submanifold P of M, and $\mathfrak{L}$ is the distribution consisting of the *vertical* subspaces, i.e., the subspaces tangent to the fibers of TM^*. All the results proven in the abstract framework will be discussed in detail for this special case in Subsection 3.3.

3.1. **The symplectic differential problem associated to a Hamiltonian system.**

It is well known that the Hamiltonian flow F_{t,t_0} consists of symplectomorphisms:

Proposition 3.1.1. *The symplectic form ω is invariant by the Hamiltonian flow F, i.e., $F_{t,t_0}^*\omega = \omega$ for all (t, t_0).* □

A sextuplet $(\mathcal{M}, \omega, H, \mathfrak{L}, \Gamma, \mathcal{P})$ where $(\mathcal{M}, \omega)$ is a symplectic manifold, H is a (time-dependent) Hamiltonian function defined on an open subset of $\mathbb{R} \times \mathcal{M}$, $\mathfrak{L}$ is a smooth distribution of Lagrangians in $\mathcal{M}$, $\Gamma : [a, b] \to \mathcal{M}$ is an integral curve of $\vec{H}$ and $\mathcal{P}$ is a Lagrangian submanifold of $\mathcal{M}$ with $\Gamma(a) \in \mathcal{P}$, will be called *a set of data for the Hamiltonian problem*.

We give some more basic definitions.

Definition 3.1.2. A vector field ρ along Γ in $\mathcal{M}$ is said to be a *solution for the linearized Hamilton (LinH) equations* if it satisfies:

$$\rho(t) = \mathrm{d}F_{t,a}(\Gamma(a))\, \rho(a). \tag{3.2}$$

We also say that ρ is a $\mathcal{P}$-*solution* for the (LinH) equations if in addition it satisfies $\rho(a) \in T_{\Gamma(a)}\mathcal{P}$.

The solutions of the (LinH) equations are precisely the variational vector fields along Γ corresponding to variations of Γ by integral curves of $\vec{H}$; the $\mathcal{P}$-solutions correspond to variations by integral curves starting on $\mathcal{P}$. It is easy to see that the set of solutions of the (LinH) equations is a vector space of dimension $2n$ and that the $\mathcal{P}$-solutions form an n-dimensional vector subspace.

Definition 3.1.3. A point $\Gamma(t)$, $t \in]a, b]$ is said to be a *$\mathcal{P}$-focal point* along Γ if there exists a non zero $\mathcal{P}$-solution ρ for the (LinH) equations such that $\rho(t) \in \mathfrak{L}_{\Gamma(t)}$. The *multiplicity* of a $\mathcal{P}$-focal point $\Gamma(t)$ is the dimension of the vector space of such ρ's.

We will now describe how, under suitable nondegeneracy hypotheses, one associates to the set of data for the Hamiltonian problem $(\mathcal{M}, \omega, H, \mathfrak{L}, \Gamma, \mathcal{P})$ an isomorphism class of sets of data for the symplectic differential problem.

Definition 3.1.4. A *symplectic $\mathfrak{L}$-trivialization* of $T\mathcal{M}$ along Γ is a smooth family of symplectomorphisms $\phi(t) : I\!R^n \oplus I\!R^{n*} \to T_{\Gamma(t)}\mathcal{M}$ such that, for all $t \in [a, b]$, $\phi(t)(L_0) = \mathfrak{L}_{\Gamma(t)}$, where $L_0 = \{0\} \oplus I\!R^{n*}$.

The existence of symplectic $\mathfrak{L}$-trivializations along Γ is easily established with elementary arguments, using the fact that $T\mathcal{M}$ restricts to a trivial vector bundle along Γ.

We will be interested also in the quotient bundle $T\mathcal{M}/\mathfrak{L}$ and its dual bundle. We have an obvious canonical identification of the dual $(T\mathcal{M}/\mathfrak{L})^*$ with the annihilator $\mathfrak{L}^o \subset T\mathcal{M}^*$; moreover, using the symplectic form, we will identify $\mathfrak{L}^o$ with $\mathfrak{L}$ by the isomorphism:

$$T_m\mathcal{M} \ni \rho \mapsto \omega(\cdot, \rho) \in T_m\mathcal{M}^*, \quad m \in \mathcal{M}. \tag{3.3}$$

A symplectic $\mathfrak{L}$-trivialization ϕ induces a trivialization of the quotient bundle $T\mathcal{M}/\mathfrak{L}$ along Γ, namely, for each $t \in [a, b]$ we define an isomorphism $\mathcal{Z}_t : I\!R^n \to T_{\Gamma(t)}\mathcal{M}/\mathfrak{L}_{\Gamma(t)}$:

$$\mathcal{Z}_t(x) = \phi(t)(x, 0) + \mathfrak{L}_{\Gamma(t)}, \quad x \in I\!R^n. \tag{3.4}$$

Given a symplectic $\mathfrak{L}$-trivialization ϕ of $T\mathcal{M}$ along Γ, we define a smooth curve $\Psi : [a, b] \to \mathrm{Sp}(2n, I\!R)$ by:

$$\Psi(t) = \phi(t)^{-1} \circ \mathrm{d}F_{t,a}(\Gamma(a)) \circ \phi(a). \tag{3.5}$$

The fact that $\Psi(t)$ is a symplectomorphism follows from Proposition 3.1.1.

We now define a smooth curve $X : [a, b] \to \mathrm{sp}(2n, I\!R)$ by setting:

$$X(t) = \Psi'(t)\Psi(t)^{-1}; \tag{3.6}$$

The $n \times n$ blocks of the matrix X will be denoted by A, B and C, as in formula (2.3). Finally, we define a Lagrangian subspace ℓ_0 of $I\!R^n \oplus I\!R^{n*}$ by:

$$\ell_0 = \phi(a)^{-1}(T_{\Gamma(a)}\mathcal{P}). \tag{3.7}$$

Suppose now that we are given another symplectic $\mathfrak{L}$-trivialization $\tilde{\phi}$ of $T\mathcal{M}$ along γ. Denote by $\tilde{\mathcal{Z}}$ the relative trivialization of $T\mathcal{M}/\mathfrak{L}$, and by $\tilde{\Psi}, \tilde{X}$ and $\tilde{\ell}_0$ the objects defined in (3.5), (3.6) and (3.7) relatively to $\tilde{\phi}$.

If we define the smooth curve $\phi_0 : [a, b] \to \mathrm{Sp}(2n, I\!R; L_0)$ by:

$$\phi_0(t) = \tilde{\phi}(t)^{-1} \circ \phi(t),$$

then obviously Ψ and $\tilde{\Psi}$ are related as in formula (2.75), and ℓ_0 is related with $\tilde{\ell}_0$ as in (2.76). Differentiating (2.75), we obtain easily (2.74). Moreover, if we set:

$$Z_t = \tilde{\mathcal{Z}}_t^{-1} \circ \mathcal{Z}_t,$$

then, $Z(t)$ is the upper-left $n \times n$ block of $\phi_0(t)$, as in (2.73).

We are ready for the following:

Definition 3.1.5. The *canonical bilinear form* of the set of data $(\mathcal{M}, \omega, H, \mathfrak{L}, \Gamma, \mathcal{P})$ is a family of symmetric bilinear forms $H_{\mathfrak{L}}(t)$ on $(T_{\Gamma(t)}\mathcal{M}/\mathfrak{L}_{\Gamma(t)})^* \simeq \mathfrak{L}^o_{\Gamma(t)} \simeq \mathfrak{L}_{\Gamma(t)}$ given by:

$$H_{\mathfrak{L}}(t) = \mathcal{Z}_t \circ B(t) \circ \mathcal{Z}_t^*, \tag{3.8}$$

where $\mathcal{Z}$ is the trivialization of $T\mathcal{M}/\mathfrak{L}$ relative to some symplectic $\mathfrak{L}$-trivialization ϕ of $T\mathcal{M}$ and B is the upper-right $n \times n$ block of the map X in (3.6). Observe that, by (2.77) and the construction of the map $\mathcal{Z}$, the right hand side of (3.8) does *not* depend on the choice of the symplectic $\mathfrak{L}$-trivialization of $T\mathcal{M}$.

We say that the set of data $(\mathcal{M}, \omega, H, \mathfrak{L}, \Gamma, \mathcal{P})$ is *nondegenerate* if $H_{\mathfrak{L}}(t)$ is nondegenerate for all $t \in [a, b]$. In this case, we can also define the symmetric bilinear form $H_{\mathfrak{L}}(t)^{-1}$ on $T_{\Gamma(t)}\mathcal{M}/\mathfrak{L}_{\Gamma(t)}$.

Given a nondegenerate set of data $(\mathcal{M}, \omega, H, \mathfrak{L}, \Gamma, \mathcal{P})$, let us consider the pair (X, ℓ_0) defined by (3.6) and (3.7). It is easily seen that the submanifold $\mathcal{P}$ and the space P defined by ℓ_0 as in (2.11) are related by the following:

$$\mathcal{Z}_a(P) = \pi(T_{\Gamma(a)}\mathcal{P}),$$

where $\pi : T_{\Gamma(a)}\mathcal{M} \to T_{\Gamma(a)}\mathcal{M}/\mathfrak{L}_{\Gamma(a)}$ is the quotient map. We set:

$$\mathcal{P}_0 = \pi(T_{\Gamma(a)}\mathcal{P}). \tag{3.9}$$

In analogy with Assumption 2.1.2, we make the following

Assumption 3.1.6. *We assume that the symmetric bilinear form $H_{\mathfrak{L}}(a)^{-1}$ is nondegenerate on $\mathcal{P}_0$. This is equivalent to requiring that $H_{\mathfrak{L}}(a)$ is nondegenerate on the annihilator $(T_{\Gamma(a)}\mathcal{P} + \mathfrak{L}_a)^o \subset \mathfrak{L}^o_a \subset T_{\Gamma(a)}\mathcal{M}^*$.*

Given a nondegenerate set of data $(\mathcal{M}, \omega, H, \mathfrak{L}, \Gamma, \mathcal{P})$ satisfying Assumption 3.1.6, the pair (X, ℓ_0) defined by (3.6) and (3.7) is a set of data for the symplectic differential problem. We say that (X, ℓ_0) is the pair *associated* to $(\mathcal{M}, \omega, H, \mathfrak{L}, \Gamma, \mathcal{P})$ by the choice of the symplectic $\mathfrak{L}$-trivialization ϕ of $T\mathcal{M}$ along Γ. We have proven above that pairs (X, ℓ_0) and $(\tilde{X}, \tilde{\ell}_0)$ associated to $(\mathcal{M}, \omega, H, \mathfrak{L}, \Gamma, \mathcal{P})$ by different choices of symplectic $\mathfrak{L}$-trivializations are isomorphic.

To define the signature of a $\mathcal{P}$-focal point along Γ, we need to introduce the following space, in analogy with formula (2.13):

$$\mathfrak{V}[t] = \big\{\rho(t) : \rho \text{ is a } \mathcal{P}\text{-solution of the (LinH) equation}\big\} \cap \mathfrak{L}_{\Gamma(t)}, \quad t \in [a, b]. \tag{3.10}$$

Using the isomorphism (3.3), it is easy to see that $\mathfrak{V}[a]$ is identified with the annihilator $(T_{\Gamma(a)}\mathcal{P} + \mathfrak{L}_a)^o$. It is easily seen that a point $\Gamma(t)$ is $\mathcal{P}$-focal if and only if $\mathfrak{V}[t]$ is not zero and that the dimension of $\mathfrak{V}[t]$ is precisely the multiplicity of $\Gamma(t)$.

Definition 3.1.7. Let $\Gamma(t)$ be a $\mathcal{P}$-focal point along Γ. The *signature* $sgn(\Gamma(t))$ is the signature of the restriction of $H_{\mathfrak{L}}(t)$ to $\mathfrak{V}[t] \subset \mathfrak{L}_t \simeq \mathfrak{L}^o_t$. $\Gamma(t)$ is said to be a nondegenerate $\mathcal{P}$-focal point if such restriction is nondegenerate. If Γ has only a finite number of $\mathcal{P}$-focal points, we define the *focal index* $\mathrm{i}_{\mathrm{foc}}(\Gamma)$ as:

$$\mathrm{i}_{\mathrm{foc}}(\Gamma) = \sum_{t \in \,]a,b]} sgn(\Gamma(t)). \tag{3.11}$$

3.2. **The index theorem for Hamiltonian systems.** Using the identification with isomorphism classes of symplectic problems, we now translate the result of the Index Theorem 2.7.10 for nondegenerate sextuplets $(\mathcal{M}, \omega, H, \mathfrak{L}, \Gamma, \mathcal{P})$. The sextuplets considered henceforth will be assumed to be nondegenerate and satisfying Assumption 3.1.6.

Recalling that the Maslov index of isomorphic symplectic problems are equal (Proposition 2.10.2), we can give the following:

Definition 3.2.1. Given a set of data $(\mathcal{M}, \omega, H, \mathfrak{L}, \Gamma, \mathcal{P})$ such that $\Gamma(b)$ is not a $\mathcal{P}$-focal point, we define its *Maslov index* $\mathrm{i}_{\mathrm{maslov}}(\Gamma)$ as the Maslov index of any pair (X, ℓ_0) associated to it by a symplectic $\mathfrak{L}$-trivialization of $T\mathcal{M}$ along Γ.

Recalling Proposition 2.10.3, we can also define an index form associated to a Hamiltonian problem as follows.

Definition 3.2.2. Let $\mathcal{H}_\Gamma$ denote the Hilbert space of sections $\mathfrak{v}$ of Sobolev class H^1 of the quotient bundle $T\mathcal{M}/\mathfrak{L}$ along Γ such that $\mathfrak{v}(a) \in \mathcal{P}_0$ and $\mathfrak{v}(b) = 0$. The *index form* I_Γ of the Hamiltonian problem is the bounded symmetric bilinear form on $\mathcal{H}_\Gamma$ defined by:

$$I_\Gamma(\mathfrak{v}, \mathfrak{w}) = I(\mathcal{Z}^{-1}\mathfrak{v}, \mathcal{Z}^{-1}\mathfrak{w}), \tag{3.12}$$

where $\mathcal{Z}$ is the trivialization of $T\mathcal{M}/\mathfrak{L}$ along Γ associated to a symplectic $\mathfrak{L}$-trivialization ϕ of $T\mathcal{M}$ as in (3.4), and I is the index form of the pair (X, ℓ_0) associated to such trivialization.

Setting $k = n_-(H_{\mathfrak{L}})$, we will now consider a smooth family of k-dimensional subspaces $\mathcal{D}_t \subset T_{\Gamma(t)}\mathcal{M}/\mathfrak{L}_{\Gamma(t)}$, $t \in [a, b]$ (this is a subbundle of the pull-back $\Gamma^*(T\mathcal{M}/\mathfrak{L})$). We assume that $H_{\mathfrak{L}}(t)^{-1}$ is negative definite on $\mathcal{D}_t$ for all t. We call such a family a *maximal negative distribution along* Γ. Given the distribution $\mathcal{D}$, we define the closed subspaces $\mathcal{S}_\mathcal{D}, \mathcal{K}_\mathcal{D} \subset \mathcal{H}_\Gamma$ by:

$$\mathcal{S}_\mathcal{D} = \big\{\mathfrak{v} \in \mathcal{H}_\Gamma : \mathfrak{v}(a) = \mathfrak{v}(b) = 0,\ \mathfrak{v}(t) \in \mathcal{D}_t,\ \forall\, t\big\}, \tag{3.13}$$

$$\mathcal{K}_\mathcal{D} = \big\{\mathfrak{v} \in \mathcal{H}_\Gamma : \mathcal{Z}^{-1}(\mathfrak{v}) \in \mathcal{K}\big\}, \tag{3.14}$$

where $\mathcal{Z}$ is the trivialization of $T\mathcal{M}/\mathfrak{L}$ along Γ associated to a symplectic $\mathfrak{L}$-trivialization ϕ of $T\mathcal{M}$ as in (3.4), and $\mathcal{K}$ is defined as in (2.34) corresponding to the distribution $\mathcal{Z}^{-1}(\mathcal{D})$ of subspaces of $\mathbb{R}^n$. By Proposition 2.10.3, the definition of the space $\mathcal{K}_\mathcal{D}$ does not depend on the choice of the trivialization ϕ.

Theorem 3.2.3. *Let $(\mathcal{M}, \omega, H, \mathfrak{L}, \Gamma, \mathcal{P})$ be a nondegenerate set of data for the Hamiltonian problem satisfying Assumption 3.1.6 and such that $\Gamma(b)$ is not a $\mathcal{P}$-focal point. Let $\mathcal{D}$ be a maximal negative distribution along Γ, and suppose that the index form I_Γ is* negative definite *on the space $\mathcal{S}_\mathcal{D}$ defined in* (3.13)*. Let's denote by $\mathcal{S}_\mathcal{D}^\perp$ the I_Γ-orthogonal complement of $\mathcal{S}_\mathcal{D}$ in $\mathcal{H}_\Gamma$, i.e.,*

$$\mathcal{S}_\mathcal{D}^\perp = \big\{\mathfrak{v} \in \mathcal{H}_\Gamma : I_\Gamma(\mathfrak{v}, \mathfrak{w}) = 0,\ \ \forall\, \mathfrak{w} \in \mathcal{S}_\mathcal{D}\big\}. \tag{3.15}$$

Then, $\mathcal{H}_\Gamma = \mathcal{S}_\mathcal{D}^\perp \oplus \mathcal{S}_\mathcal{D}$, $\mathcal{S}_\mathcal{D}^\perp = \mathcal{K}_\mathcal{D}$ and:

$$n_-(I_\Gamma|_{\mathcal{K}_\mathcal{D}}) = n_-(H_{\mathfrak{L}}(a)^{-1}|_{\mathcal{P}_0}) + \mathrm{i}_{\mathrm{maslov}}(\Gamma). \tag{3.16}$$

Proof. Use a symplectic $\mathfrak{L}$-trivialization of $T\mathcal{M}$ along Γ to associate to the given sextuplet $(\mathcal{M}, \omega, H, \mathfrak{L}, \Gamma, \mathcal{P})$ a set of data for the symplectic differential problem, for which the results of Section 2 apply. The thesis follows from Corollary 2.6.10, Lemma 2.6.11, Theorem 2.7.10 and Corollary 2.10.7. $\square$

3.3. **The case of a hyper-regular Hamiltonian system.** We will now consider the special case of a Hamiltonian problem defined in the cotangent bundle $\mathcal{M} = TM^*$ of a smooth n-dimensional manifold M endowed with the canonical symplectic form ω. Recall that ω is defined by $\omega = -\mathrm{d}\theta$, where θ is the canonical 1-form on TM^*, given by $\theta_p(\rho) = p(\mathrm{d}\pi_p(\rho))$ where $\pi : TM^* \to M$ is the projection, $p \in TM^*$, $\rho \in T_pTM^*$.

We will prove that, in this situation, if the Hamiltonian is hyper-regular, i.e., if it corresponds to a Lagrangian, then the index form I_Γ is the second variation of the Lagrangian action principle. In the general context of subsection 3.1 we do not know whether there exists an interpretation of I_Γ as a second variation of some functional.

A local chart $(q_1, \ldots, q_n)$ on M induces a local chart $(q_1, \ldots, q_n, p_1, \ldots, p_n)$ on TM^*, which is symplectic. We consider the Lagrangian distribution $\mathfrak{L}$ on TM^* given by:

$$\mathfrak{L}_p = \mathrm{Ker}(\mathrm{d}\pi_p) = T_p(T_{\pi(p)}M^*) \simeq T_{\pi(p)}M^*, \quad p \in TM^*. \tag{3.17}$$

We also call $\mathfrak{L}_p$ the *vertical subspace* of T_pTM^*.

Let P be a smooth submanifold of M; then the annihilator

$$\mathcal{P} = TP^o = \big\{ p \in TM^* : \pi(p) = m \in P \text{ and } p|_{T_mP} = 0 \big\}$$

is a Lagrangian submanifold of TM^*. Indeed, θ vanishes on TP^o and so does ω.

We consider a Hamiltonian H defined on an open subset of $I\!R \times TM^*$ and an integral curve $\Gamma : [a,b] \to TM^*$ of $\vec{H}$ such that $\Gamma(a) \in TP^o$; let γ be the projection in M of Γ:

$$\gamma = \pi \circ \Gamma,$$

so that $\gamma(a) \in P$.

Using the projection π, we can identify the quotient $T_p\mathcal{M}/\mathfrak{L}_p$ with the tangent space $T_{\pi(p)}M$, and the space $\mathcal{P}_0$ defined in (3.9) becomes identified with $T_{\gamma(a)}P$. By these identifications, we can describe the Hilbert space $\mathcal{H}_\Gamma$ of Definition 3.2.2 as the space of vector fields $\mathfrak{v}$ along γ in M, of Sobolev class H^1, such that $\mathfrak{v}(a) \in T_{\gamma(a)}P$ and $\mathfrak{v}(b) = 0$.

Let $q = (q_1, \ldots, q_n)$ be a local chart in M and let $(q,p) = (q_1, \ldots, q_n, p_1, \ldots, p_n)$ be the corresponding chart in TM^*; suppose that we have a subinterval $[c,d] \subset [a,b]$ such that $\gamma([c,d])$ is contained in the domain of the chart q. For each $t \in [c,d]$, the differential of the chart (q,p) gives a symplectomorphism from $I\!R^n \oplus I\!R^{n*}$ to $T_{\Gamma(t)}TM^*$ which takes $L_0 = \{0\} \oplus I\!R^{n*}$ to the vertical space $\mathfrak{L}_{\Gamma(t)}$. Suppose that ϕ is a symplectic $\mathfrak{L}$-trivialization of TTM^* along Γ such that, for $t \in [c,d]$, $\phi(t)$ is the symplectomorphism induced by the chart (q,p). In this case we say that the symplectic $\mathfrak{L}$-trivialization is *compatible* with the chart (q,p) on the interval $[c,d]$; the corresponding trivialization $\mathcal{Z}$ of the quotient $T\mathcal{M}/\mathfrak{L}$ in the interval $[c,d]$ is given by the differential of the chart q. Now, consider the pair (X, ℓ_0) of data for the symplectic differential problem associated to ϕ; for $t \in [c,d]$, we have:

$$X = \begin{pmatrix} \frac{\partial^2 H}{\partial q \partial p} & \frac{\partial^2 H}{\partial p^2} \\ -\frac{\partial^2 H}{\partial q^2} & -\frac{\partial^2 H}{\partial p \partial q} \end{pmatrix}. \tag{3.18}$$

This is obtained by observing that solutions (v, α) of the symplectic differential system with coefficient matrix (3.18) correspond by ϕ to solutions ρ of the linearized Hamilton equations; such ρ's are variational vector fields corresponding to variations of Γ by integral curves of $\vec{H}$. In this setup, the canonical bilinear form $H_\mathfrak{L}(t)$ of (3.8) is a bilinear form in $T_{\gamma(t)}M^*$, and we have the following:

Proposition 3.3.1. *The canonical bilinear form $H_\mathfrak{L}(t)$ equals the Hessian at $\Gamma(t)$ of the map $p \mapsto H(t,p)$, defined in a neighborhood of $\Gamma(t)$ in the fiber $T_{\gamma(t)}M^*$.*

Proof. It follows directly from (3.18). □

Let us now assume that H is a *hyper-regular* Hamiltonian on an open subset $U \subset \mathbb{R} \times TM^*$, i.e., its fiber derivative is a (global) diffeomorphism between U and an open subset $V \subset \mathbb{R} \times TM$ (see [1]); let L be the corresponding hyper-regular Lagrangian on V. The map L defines an *action functional* on the set of C^1-curves $\vartheta : [a,b] \to M$ such that $(t, \vartheta'(t)) \in V$ for all $t \in [a,b]$, by:

$$\vartheta \longmapsto \int_a^b L(t, \vartheta'(t)) \, \mathrm{d}t. \tag{3.19}$$

The set of such curves ϑ has the structure of an infinite dimensional Banach manifold, and the action functional above is smooth. It is well known (see [1, Chapter 3]) that the critical points of the restriction of the action functional (3.19) to the space of C^1-curves connecting the submanifold P and the point $\gamma(b)$ are the curves $\vartheta : [a,b] \to M$ such that $\vartheta(b) = \gamma(b)$ and $t \mapsto \mathrm{d}\big(L|_{(\{t\}\times T_{\vartheta(t)}M)\cap V}\big)(t, \vartheta'(t))$ is an integral curve of $\vec{H}$ starting at $T\mathcal{P}_0$.

Hence, γ is a critical point of the restriction of the action functional, and the Hessian of such functional at γ defines a bounded symmetric bilinear form on the Banach space of C^1-vector fields $\mathfrak{v}$ along γ such that $\mathfrak{v}(a) \in TP$ and $\mathfrak{v}(b) = 0$. This Hessian is computed in the following:

Proposition 3.3.2. *The Hessian of the action functional* (3.19) *at* γ *has a (unique) extension to a bounded symmetric bilinear form on the Hilbert space* $\mathcal{H}_\Gamma$*, which is given by* I_Γ *(see Definition 3.2.2).*

Proof. If $\mathfrak{v}, \mathfrak{w}$ are vector fields along γ with "small" support, then the computation of $I_\Gamma(\mathfrak{v}, \mathfrak{w})$ can be done in local coordinates using (3.18). In this case, the conclusion is easily obtained. For the general case, one simply observes that both the index form and the Hessian are bilinear, and that any smooth vector field along γ can be written as a finite sum of vector fields with "small" support. □

3.4. **An interpretation of the results using connections.** In order to give a geometrical, and perhaps more intuitive, idea of the results presented, we will describe the theory of Subsection 3.3 in terms of a torsion-free linear connection on TM.

Let us consider the setup described at the beginning of Subsection 3.3; in addition, we will consider an arbitrary torsion–free (i.e., symmetric) connection ∇ on the tangent bundle TM. We consider the *curvature tensor* R of ∇ chosen with the following sign convention: $R(X,Y) = \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]}$; the connection ∇ induces a connection ∇^* in TM^*. For each $p \in TM^*$, the connection ∇^* produces a direct sum decomposition of T_pTM^* into a *horizontal* and a *vertical* space in a standard way. The vertical space is identified with the fiber $T_{\pi(p)}M^*$, and, using $\mathrm{d}\pi(p)$, the horizontal space of T_pTM^* is identified with the tangent space $T_{\pi(p)}M$. These identifications give isomorphisms:

$$TTM^* \simeq \pi^*TM \oplus \pi^*TM^*, \quad T^*TM^* \simeq \pi^*TM^* \oplus \pi^*TM, \tag{3.20}$$

where π^* denotes the pull-back of fiber bundles. Hence, we have connections $\pi^*\nabla \oplus \pi^*\nabla^*$ and $\pi^*\nabla^* \oplus \pi^*\nabla$ on TTM^* and T^*TM^* respectively induced by the above isomorphisms. Now, given the Hamiltonian function H, for each $t \in \mathbb{R}$ we have a smooth function H_t in (an open subset of) TM^*, hence $\mathrm{d}H_t$ is a smooth section of T^*TM^*. Using the decomposition (3.20), we write $\mathrm{d}H = \left(\frac{\partial H}{\partial q}, \frac{\partial H}{\partial p}\right)$, where $\frac{\partial H}{\partial q}$ and $\frac{\partial H}{\partial p}$ denote the restrictions of $\mathrm{d}H$ to the horizontal and the vertical subspaces of TTM^* respectively. More in general, we will denote by $\frac{\partial}{\partial q}$ and $\frac{\partial}{\partial p}$ respectively the restrictions to the horizontal and the vertical

subspaces of TTM^* of covariant derivatives. One should observe that the connection in TTM^* is not torsion-free; indeed its torsion is easily computed as:

$$T_p(\rho_1,\rho_2) = (0, R(\mathrm{d}\pi(\rho_2), \mathrm{d}\pi(\rho_1))^* p),$$

for $p \in TM^*$ and $\rho_1,\rho_2 \in T_pTM^*$.

Using the decomposition (3.20) of TTM^*, from the Cartan's identity for exterior differentiation and the symmetry of the connection in TM, we have the following identity for the canonical symplectic form of TM^*:

$$\omega(\rho_1,\rho_2) = [\rho_2]_{\mathrm{vert}}(\mathrm{d}\pi(\rho_1)) - [\rho_1]_{\mathrm{vert}}(\mathrm{d}\pi(\rho_2)), \tag{3.21}$$

for $p \in TM*$, $\rho_1,\rho_2 \in T_pTM^*$, where $[\cdot]_{\mathrm{vert}}$ denotes projection onto the vertical space. Recalling that Γ is an integral curve of $\vec{H}$ in TM^* and that $\gamma = \pi\circ\Gamma$ is its projection in M, the Hamilton equations are written as:

$$\frac{\mathrm{d}}{\mathrm{d}t}\gamma(t) = \frac{\partial H_t}{\partial p}(\Gamma(t)); \quad \frac{\mathrm{D}}{\mathrm{d}t}\Gamma(t) = -\frac{\partial H_t}{\partial q}(\Gamma(t)), \tag{3.22}$$

where $\frac{\mathrm{d}}{\mathrm{d}t}$ denotes standard derivative and $\frac{\mathrm{D}}{\mathrm{d}t}$ denotes covariant derivative. Considering a smooth variation $\Gamma(t,s)$ of Γ, $s \in]-\varepsilon,\varepsilon[$, by integral curves of $\vec{H}$, with $\Gamma(t,0) = \Gamma(t)$ and with variational vector field $\rho(t) = \frac{\mathrm{d}}{\mathrm{d}s}\Gamma(t,0)$, from (3.22) we compute the linearized Hamilton equations:

$$\frac{\mathrm{D}}{\mathrm{d}t}v = \frac{\partial^2 H_t}{\partial q\partial p}v + \frac{\partial^2 H_t}{\partial p^2}\alpha; \quad \frac{\mathrm{D}}{\mathrm{d}t}\alpha = -\frac{\partial^2 H_t}{\partial q^2}v - R\left(\frac{\mathrm{d}}{\mathrm{d}t}\gamma, v\right)^*\Gamma - \frac{\partial^2 H_t}{\partial p\,\partial q}\alpha, \tag{3.23}$$

where $\rho = (v,\alpha)$ is a smooth vector field along Γ. We now make an appropriate choice of a symplectic $\mathfrak{L}$-trivialization of TTM^* along Γ; we consider a parallel trivialization $\{v_1,\ldots,v_n\}$ of TM along γ, and let $\{\alpha_1,\ldots,\alpha_n\}$ be the corresponding dual trivialization of TM^* along γ, which is also parallel. Using the decomposition of TTM^* in (3.20) and formula (3.21), it is easily seen that $\{v_1,\ldots,v_n,\alpha_1,\ldots,\alpha_n\}$ gives a symplectic $\mathfrak{L}$-trivialization ϕ of TTM^* along Γ. The coefficients of the symplectic differential system associated to the Hamiltonian problem by the trivialization ϕ are the coordinates in the basis $\{v_1,\ldots,v_n\}$ of the coefficients of the system (3.23). Identifying tensors in TM^* with their matrices in the basis $\{v_1.\ldots,v_n,\alpha_1,\ldots,\alpha_n\}$, we have:

$$A = \frac{\partial^2 H_t}{\partial q\partial p}, \quad B = \frac{\partial^2 H_t}{\partial p^2}, \quad C = -\frac{\partial^2 H_t}{\partial q^2} - \Gamma\circ R\left(\frac{\mathrm{d}}{\mathrm{d}t}\gamma,\cdot\right). \tag{3.24}$$

Using this machinery, we are now able to give a geometrical description of the space $\mathcal{K}_{\mathcal{D}}$ introduced in (3.14). Let's assume that $\mathcal{D}$ is a smooth distribution in M defined in a neighborhood of the image of the curve γ, such that $\mathcal{D}_{\gamma(t)} = \mathcal{D}_t$ for all t. We say that a curve $\vartheta : [a,b] \to M$ is a *solution of the Hamilton equations in the direction $\mathcal{D}$* if there exists a curve $\Theta : [a,b] \to TM^*$, with $\pi\circ\Theta = \vartheta$, such that the following system is satisfied:

$$\frac{\mathrm{d}}{\mathrm{d}t}\vartheta(t) = \frac{\partial H_t}{\partial p}(\Theta(t)); \quad \frac{\mathrm{D}}{\mathrm{d}t}\Theta(t)\big|_{\mathcal{D}(\vartheta(t))} = -\frac{\partial H_t}{\partial q}(\Theta(t))\big|_{\mathcal{D}(\vartheta(t))}. \tag{3.25}$$

An easy computation gives the proof of the following:

Proposition 3.4.1. *A vector field $\mathfrak{v} \in \mathcal{H}_\Gamma$ of Sobolev class H^2 is in $\mathcal{K}_{\mathcal{D}}$ if and only if it is the variational vector field of a variation of γ by solutions of the Hamilton equations in the direction of $\mathcal{D}$ connecting P and $\gamma(b)$.* □

3.5. **The Morse Index Theorem in semi-Riemannian Geometry.** We will now use the results of the previous sections to study the case of a *geodesic Hamiltonian* on a semi-Riemannian manifold (M, g), i.e., $H : TM^* \to I\!R$ is given by $H(p) = \frac{1}{2} g^{-1}(p, p)$ where g^{-1} denotes the inner product on TM^* induced by g. Recall that a semi-Riemannian metric on a manifold M is a smooth symmetric $(2, 0)$-tensor field g which is nondegenerate at every point. It is easy to see that H is hyper-regular, the corresponding Lagrangian $L : TM \to I\!R$ is given by $L(v) = \frac{1}{2} g(v, v)$, and the action functional is given by:

$$\vartheta \longmapsto \frac{1}{2} \int_a^b g(\vartheta', \vartheta') \, \mathrm{d}t.$$

Let ∇ be the covariant derivative of the Levi–Civita connection of g; in the notation of Subsection 3.4, we have:

$$\frac{\partial H}{\partial q} = 0, \quad \frac{\partial H}{\partial p} = g^{-1}(p), \quad \frac{\partial^2 H}{\partial q^2} = 0, \quad \frac{\partial^2 H}{\partial q \partial p} = 0, \quad \frac{\partial^2 H}{\partial p \partial q} = 0, \quad \frac{\partial^2 H}{\partial p^2} = g^{-1}.$$

By (3.22), a solution of the Hamilton equations is a curve Γ of the form $\Gamma = g(\gamma')$, where γ is a geodesic in (M, g). As in Subsection 3.4, using a parallel trivialization of TM along γ, we obtain an associated symplectic differential problem whose coefficients (3.24) are given by:

$$A = 0, \quad B = g^{-1}, \quad C = g(R(\gamma', \cdot)\, \gamma', \cdot). \tag{3.26}$$

Let P be a submanifold of M; the condition that $\Gamma(a) \in TP^o$ means that the underlying geodesic γ starts orthogonally to P. Assumption 3.1.6 in this context means that g is nondegenerate at $T_{\gamma(a)}P$; by (3.23), a solution v of the linearized Hamilton equations is a *Jacobi field* along γ, i.e.,

$$v'' = R(\gamma', v)\, \gamma' \quad \text{and} \quad \alpha_v = g(v', \cdot).$$

The $\mathcal{P}$-solutions are the P-Jacobi fields along γ, which are the Jacobi fields v satisfying the initial conditions:

$$v(a) \in T_{\gamma(a)}P, \quad v'(a) + S^P_{\gamma'(a)}(v(a)) \in T_{\gamma(a)}P^\perp,$$

where $S^P_{\gamma'(a)}$ is the second fundamental form of P at $\gamma(a)$ in the normal direction $\gamma'(a)$.

A P-focal point along γ is a point $\gamma(t_0)$, $t_0 \in \,]a, b]$, such that there exists a non zero P-Jacobi field along γ vanishing at t_0. For each $t \in \,]a, b]$, we set:

$$\mathcal{J}[t] = \big\{ v(t) : v \text{ is a } P\text{-Jacobi field along } \gamma \big\};$$

it is an easy observation that the point $\gamma(t_0)$ is P-focal if and only if $\mathcal{J}[t] \ne T_{\gamma(t_0)}M$. The *multiplicity* of the P-focal point $\gamma(t_0)$ is the codimension of $\mathcal{J}[t_0]$; $\gamma(t_0)$ is nondegenerate if g is nondegenerate on $\mathcal{J}[t_0]$ and the signature of $\gamma(t_0)$ is the signature of the restriction of g to $\mathcal{J}[t_0]^\perp$, where $\perp$ denotes the orthogonal complement with respect to g.

If $\gamma(b)$ is not P-focal, the Maslov index $\mathrm{i}_{\mathrm{maslov}}(\gamma)$ is defined as the Maslov index $\mathrm{i}_{\mathrm{maslov}}(\Gamma)$ of the solution of the Hamiltonian H; if all the P-focal points along γ are nondegenerate, then, by Theorem 2.3.3, $\mathrm{i}_{\mathrm{maslov}}(\gamma)$ is equal to the sum of the signatures of all the P-focal points along γ.

The index form I_Γ is given by:

$$I_\Gamma(v, w) = \int_a^b \Big[g(v', w') + g(R(\gamma', v)\, \gamma', w) \Big] \, \mathrm{d}t - g(S^P_{\gamma'(a)} v(a), w(a)). \tag{3.27}$$

Let $k = n_-(g) = n_-(\frac{\partial^2 H}{\partial p^2}) = n_-(H_{\mathfrak{L}})$; in order to apply Theorem 3.2.3 we need to determine a k-dimensional distribution along γ consisting of subspaces where the metric

tensor g is negative definite. Let $\mathcal{D}$ be one such distribution, generated by the vector fields $Y_1, \ldots, Y_k$ along γ. The space $\mathcal{K}_{\mathcal{D}}$ defined in (3.14) can be described as:

$$\begin{aligned}\mathcal{K}_{\mathcal{D}} = \big\{ v \in \mathcal{H}_\Gamma : g(v', Y_i) \in H^1([a,b]; I\!R) \text{ and} \\ g(v', Y_i)' = g(v', Y_i') + g(R(\gamma', v)\,\gamma', Y_i),\ \forall\, i = 1, \ldots, k \big\}.\end{aligned} \tag{3.28}$$

The matrices $\mathcal{B}$, $\mathcal{C}$ and $\mathcal{I}$ appearing in the reduced symplectic system (2.36) are given by:

$$\mathcal{B}_{ij} = g(Y_i, Y_j), \quad \mathcal{C}_{ij} = g(Y_j', Y_i), \quad \mathcal{I}_{ij} = g(Y_i', Y_j') + g(R(\gamma', Y_i)\,\gamma', Y_j). \tag{3.29}$$

We can now restate Theorem 3.2.3 in the form of a Morse Index Theorem for semi-Riemannian geometry:

Theorem 3.5.1. *Let (M, g) be a semi-Riemannian manifold with $n_-(g) = k$, let $P \subset M$ be a smooth submanifold and $\gamma : [a, b] \to M$ be a geodesic with $\gamma'(a) \in TP^\perp$. Suppose that g is nondegenerate on $T_{\gamma(a)}P$ and that $\gamma(b)$ is not P-focal. Let $Y_1, \ldots, Y_k$ be a family of smooth vector fields along γ generating a k-dimensional distribution $\mathcal{D}$ along γ on which g is negative definite. If the reduced symplectic system* (2.36) *defined by the matrices* (3.29) *has no focal instants, then the index of I_Γ (*(3.27)*) on the space $\mathcal{K}_{\mathcal{D}}$ defined in* (3.28) *is given by:*

$$n_-(I_\Gamma\big|_{\mathcal{K}_{\mathcal{D}}}) = \mathrm{i}_{\mathrm{maslov}}(\gamma) + n_-(g\big|_{T_{\gamma(a)}P}). \qquad \square$$

Observe that if (M, g) is Riemannian, then Theorem 3.5.1 is the Morse Index Theorem for Riemannian geodesics with variable initial points; if (M, g) is Lorentzian and γ is a timelike geodesic, then it is not hard to see that, taking $Y_1 = \gamma'$ in Theorem 3.5.1, we obtain the Timelike Morse Index Theorem in Lorentzian geometry (see [3, 23]). Theorem 3.5.1 generalizes [11, Theorem 6.1].

We present some examples where the Theorem 3.5.1 applies.

Example 3.5.2. If $Y_1, \ldots, Y_k$ are Jacobi fields such that $g(Y_i', Y_j)$ is symmetric, then the reduced symplectic system (2.36) has no focal instants (see Example 2.6.3). In this case, the space $\mathcal{K}_{\mathcal{D}}$ can be easily described as:

$$\mathcal{K}_{\mathcal{D}} = \big\{ v \in \mathcal{H}_\Gamma : g(v', Y_i) - g(v, Y_i') = \text{constant}, \ \ i = 1, \ldots, k \big\}.$$

Example 3.5.3. Let G be a k-dimensional Lie group acting on M by isometries with no fixed points, or more in general, having only discrete isotropy groups. Suppose that g is negative definite on the orbits of G. If $\gamma'(a)$ is orthogonal to the orbit of the commutator subgroup $[G, G]$ (for instance if G is abelian), then Theorem 3.5.1 can be applied by considering the distribution $\mathcal{D}$ tangent to the orbits of G. Observe indeed that $\mathcal{D}$ is generated by k linearly independent Killing vector fields $Y_1, \ldots, Y_k$ on M, which therefore restrict to Jacobi fields along any geodesic. Then, one falls into the situation of Example 3.5.2 by observing that the symmetry of $g(Y_i', Y_j)$ follows from the orthogonality with the orbits of $[G, G]$:

$$g(Y_i', Y_j) - g(Y_i, Y_j') = -g(\nabla_{Y_j} Y_i, \gamma') + g(\nabla_{Y_i} Y_j, \gamma') = g([Y_i, Y_j], \gamma') = 0.$$

In this situation, the space $\mathcal{K}_{\mathcal{D}}$ can be described as the space of variational vector fields along γ corresponding to variations of γ by curves that are *geodesics along* $\mathcal{D}$, i.e., whose second derivatives are orthogonal to $\mathcal{D}$ (see Proposition 3.4.1).

Example 3.5.4. Suppose that the bilinear form $g(R(\gamma', \cdot)\,\gamma', \cdot)$ is negative semidefinite along the geodesic γ. Then, Theorem 3.5.1 can be applied by considering any k-dimensional parallel distribution $\mathcal{D}$ along γ where g is negative definite (see Example 2.6.4). If $k = 1$, i.e., if (M, g) is Lorentzian, this is a condition on the sign of the sectional curvature of timelike planes containing γ'.

We conclude with the observation that from Theorem 2.8.1 one can easily obtain a semi-Riemannian Morse Index Theorem for geodesics starting and ending orthogonally to two submanifolds of M.

REFERENCES

[1] R. Abraham, J. E. Marsden, *Foundations of Mechanics*, 2nd Edition, Benjamin/Cummings, Advanced Book Program, Reading, Mass., 1978.

[2] V. I. Arnol'd, *Characteristic Class Entering in Quantization Conditions*, Funct. Anal. Appl. **1** (1967), 1–13.

[3] J. K. Beem, P. E. Ehrlich, K. L. Easley, *Global Lorentzian Geometry*, Marcel Dekker, Inc., New York and Basel, 1996.

[4] R. Bott, *On the Iteration of Closed Geodesics and the Sturm Intersection Theory*, Commun. Pure Appl. Math. **9** (1956), 171–206.

[5] E. A. Coddington, N. Levinson, *Theory of Ordinary Differential Equations*, McGraw–Hill Book Company, New York, Toronto, London, 1955.

[6] M. do Carmo, *Riemannian Geometry*, Birkhäuser, Boston, 1992.

[7] C. Conley, E. Zehnder, *Morse-type Index Theory for Flows and Periodic Solutions for Hamiltonian Equations*, Comm. Pure Appl. Math. **37** (1984), 207–253.

[8] J. J. Duistermaat, *On the Morse Index in Variational Calculus*, Adv. in Math. **21** (1976), 173–195.

[9] H. M. Edwards, *A Generalized Sturm Theorem*, Ann. of Math. **80** (1964), 22–57.

[10] I. Ekeland, *An Index Theory for Periodic Solutions of Convex Hamiltonian Systems*, Proc. of Symposia in Pure Math. vol. 45 (1986), Part I, 395–423.

[11] F. Giannoni, A. Masiello, P. Piccione, D. Tausk, *A Generalized Index Theorem for Morse–Sturm Systems and Applications to semi-Riemannian Geometry*, preprint 1999 (`LANL math.DG/9908056`), to appear in the Asian Journal of Mathematics.

[12] V. Guillemin, S. Sternberg, *Geometric Asymptotics*, Mathematical Surveys and Monographs n. 14, AMS, Providence RI, 1990.

[13] A. D. Helfer, *Conjugate Points on Spacelike Geodesics or Pseudo-Self-Adjoint Morse-Sturm-Liouville Systems*, Pacific J. Math. **164**, n. 2 (1994), 321–340.

[14] M. W. Hirsch, *Differential Topology*, Springer–Verlag, 1976.

[15] D. Kalish, *The Morse Index Theorem where the Ends are Submanifolds*, Trans. Am. Math. Soc. **308**, n. 1 (1988), 341–348.

[16] Y. Long, *A Maslov-type Index Theory for Symplectic Paths*, Top. Meth. Nonlin. Anal. **10** (1997), 47–78.

[17] A. Masiello, *Variational Methods in Lorentzian Geometry*, Pitman Research Notes in Mathematics **309**, Longman, London 1994.

[18] J. Mawhin, M. Willem, *Critical Point Theory and Hamiltonian Systems*, Springer–Verlag, Berlin, 1989.

[19] F. Mercuri, P. Piccione, D. Tausk, *Stability of the Focal and the Geometric Index in semi-Riemannian Geometry via the Maslov Index*, Technical Report RT-MAT 99-08, Mathematics Department, University of São Paulo, Brazil, 1999. (`LANL math.DG/9905096`)

[20] J. Milnor, *Morse Theory*, Princeton Univ. Press, Princeton, 1969.

[21] Perlick, V. *Ray optics, Fermat's principle and applications to general relativity*, Springer Lecture Notes in Physics, m61.

[22] V. Perlick, P. Piccione, D. Tausk, *A Variational Principle for Light Rays in a General Relativistic Plasma*, preprint 1999.

[23] P. Piccione, D. V. Tausk, *A Note on the Morse Index Theorem for Geodesics between Submanifolds in semi-Riemannian Geometry*, J. Math. Phys. **40**, vol. 12 (1999), 6682–6688.

[24] P. Piccione, D. V. Tausk, *The Maslov Index and a Generalized Morse Index Theorem for Non Positive Definite Metrics*, Comptes Rendus de l'Académie de Sciences de Paris, vol. 331, 5 (2000), 385–389.

[25] D. Salamon, E. Zehnder, *Morse Theory for Periodic Solutions of Hamiltonian Systems and the Maslov Index*, Commun. Pure Appl. Math. **45** (1992), 1303–1360.

[26] F. Treves, *Introduction to Pseudodifferential Operators*, Plenum, New York, 1982.

DEPARTAMENTO DE MATEMÁTICA, UNIVERSIDADE DE SÃO PAULO, BRAZIL

E-mail address: `piccione@ime.usp.br, tausk@ime.usp.br`

URL: `http://www.ime.usp.br/˜piccione, http://www.ime.usp.br/˜tausk`